

FIGURE 1

CGGCAGCCCTGACGTGATGAGCTCAACCAGCAGAGACATTCCATCCCAAGAGAGGGTCTGCGTGACGCCGCCGGGAGG
 CCACCCCTCAGCAAGACCACCGTACAGTTGGTGGAAAGGGGTGACAGCTGCATTCTCCTGTGCCCTACCACGTAACCAAA
AATGAAGGAGAACTACTGTTACAAGCCGCCCTGGTGTGCCCTGGCATGCTGTGCCACAGCCATGCCCTTGCCCCAG
 AGCGGCGGGGGCACCTGCCCTCCTCCATGGCACCATGAGAAGGGCAAGGAGGGCAGGTGCTACAGCGCTCC
 AACGCTGGCTGGCTGGAACCAGTTCTCGTGTAGAGGAGTACACCGGGCTGACCCCGTGTGGCAGGCT
 TCATTCAAGATATTGACTCTGGTGTAGGGAACATTAATAACATTCTCTCAGGGAAAGGAGGCTGGAACCATTGGTGA
 TTGATGACAAATCAGGGAACATTGATGCCACCAAGACGTTGGATCGAGAAGAGAGAGCCCAGTACACGTTGATGGCT
 CAGGCAGTGGACAGGGACACCAATCGGCCACTGGAGCCACCGTCGAATTGTCAGGTCCAGGACATTAATGA
 CAACCCCTCCGGAGTTCTGCACGAGACCTATCATGCCAACGTGCCTGAGAGGTCCAATGTGGAACGTCAGTAATCC
 AGGTGACAGCTTCAGATGCAGATGACCCACTTATGAAATAGGCCAAGTTAGTGTACAGTATCCTCGAAGGACAA
 CCCTATTTCGGTGGAAAGCACAGACAGGTATCATCAGAACAGCCCTACCCAACATGGACAGGGAGGCAAGGAGGA
 GTACCACGTGGTGTCCAGGCCAAGGACATGGTGGACATGGTGGACATGGCGGACTCTCAGGGACAACCAAAGTGCAGATCA
 CACTGACCGATGTCATGACAACCCACCAAAGTTCCGAGAGGCTATACAGATGTCAGGAGGAGGAGGAGGAGGAGGAGG
 CCTGGGGAGGAAGTAGGAAGAGTGAAAGCTAAAGATCCAGACATTGGAGAAAATGGCTTAGTCACATACAATTG
 TGATGGAGATGGTATGGAATGTTGAAATCACAAACGGACTATGAAACACAGGGGGGTGATAAGCTGAAAAAGC
 CTGTTAGATTTGAAACCGAAAGAGCCTATAGCTGAAGGTAGAGGCAGCCAACGTGCACATGACCCGAAGTTATC
 AGCAATGGCCCTTCAAGGACACTGTGACCGTCAAGATCTCAGTAGAAGATGTCAGGAGGAGGAGGAGGAGGAGG
 CCCAAGTTACATCCACGAAGTCCAAGAAAATGCAGCTGCTGGCACCGTGGTTGGAGAGTGCATGCCAAAGACCCCTG
 ATGCTGCCAACAGCCGATAAGGTATTCCATCGATCGTACACTGACCTCGACAGATTTTCACTATTAATCCAGAG
 GATGGTTTATTAAAACACAAACCTCTGGATAGAGAGGAAACAGCCTGGCTCAACATCAGTCTTGCAGCAGA
 AATCCACAATCGGCATCAGGAAGCCAAGTCCCAGTGGCATTAGGGCCTTGATGTCACAGATAATGCTCCAAGT
 TTGCTGCCCTTATGAAGGTTCATCTGTGAGAGTGATCAGACCAAGCCACTTTCAACCAGCAATTGTTACAATT
 AGTGCAGATGACAAGGATGACACGGCCAATGGACCAAGATTTATCTCAGCCTACCCCTGAAATCATTACAATCC
 AAATTCACAGTCAGAGACAACCGAGATAACACACAGCAGGGCTGTACGCCGGCGTGGAGGGTCAGTCGGCAGAAC
 AGGACTTGTACCTCTGCCCATAGTGTACAGCGATGGCGGATCCGCCATGAGTAGCACCAACACCCCTACCATC
 AAAGTCTCGGGTGCACGTGAACGGGCACTGCTCTGCACAGCAGAGGCTACATTCTGAACGCCGGCTGAG
 CACAGGCCCTGATGCCATCCTGCCATCGTCAATTCTCTGGCATTGAGTAGTATTGTTGTGACCCCTGAGAA
 GGCAAAAGAAAGAACCACTCATTGTCTTGAGGAAGAAGATGTCGTGAGAACATCATTACTATGATGATGAAGGG
 GGTGGGGAGAACACAGAACGCCCTTGATATTGCCACCCCTCCAGAACCTGATGGTATCAATGGATTATCCCCCG
 CAAAGACATCAAACCTGAGTATCAGTACATGCCATGACCTGGCTCCGCCAGGCCAACAGCGTGGATGTCGATG
 ACTTCATCAACACGAGAACAGGAGGGCAGACAATGACCCACGGCTCCTTATGACTCATTCAAATCTACGGT
 TATGAAGGCAGGGCTCAGTGGCGGGTCCCTGAGCTCCCTAGAGTCGGCCACCAAGATTGAGACTTGGACTATGA
 TTATCTACAGAACTGGGACCTCGTTTAAGAAACTAGCAGATTGTATGGTCAAAGACACTTTGATGACGATT
CTTAACAATAACGATAAAATTGGCCTTAAGAACTGTGCTGGCGTCTCAAGAATCTAGAAGATGTGTAACAGGT
 ATTTTTT

FIGURE 2

AACTCAAACCTCTCTGGGAAAACGC GGTC TTGCTCCCTCCGGAGTGGCCTTGGCAGGGTGTGGAGCCCTCG
GTCTGCCCGTCCGGTCTCTGGGGCCAAGGCTGGTTCCCTCATGTATGGCAAGAGCTACTCGTGC GG TGCTTC
TTCTCCTGGCATA CAGCTCACAGCTTTGGCTATAGCAGCTGTGGAAATTATACCTCCGGGTGCTGGAGGCT
GTTAATGGGACAGATGCTCGTTAAAATGCAC T TCTCCAGCTTGCCCCTGTGGGTGATGCTCTAACAGTGACCTG
GAATTTCGTCCTCTAGACGGGGGACCTGAGCAGTTGTTACTACCACATAGATCCCTCCAACCCATGAGTG
GGCGGTTAAAGGACCGGGTGTCTGGGATGGGAATCCTGAGCGGTACGATGCCTCCATCCTCTTGAAACTGCAG
TTCGACGACAATGGGACATA CACCTGCCAGGTGAAGAACCCACCTGATGTTGATGGGTGATAGGGGAGATCCGGCT
CAGCGTCGTGCACACTGTACGCTTCTCTGAGATCCACTTCCTGGCTCGGCCATTGGCTTGCACTGATGTA
TCATAATAGTAATTGTAGTGGCCTCTCCAGCATTACCGAAAAAGCGATGGGCCAAAGAGCTCATAAAGTGGTG
GAGATAAAATCAAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCTGTTATTTAGAAGACACAGACTAAC
ATTTAGATGGAAGCTGAGATGATTCCAAGAACAAAGAACCCTAGTATTCTGAAGTTAATGGAAACTTTCTTG
GCTTTCCAGTTGTGACCCGTTTCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGC
CAGCACAGTGCTCCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTATTAAATTTCAGAGTGTAAATT
TTTCAAGTGCTCATTAGTTTATAAACAAAGAAGCTACATTTGCCCTTAAGACACTACTACAGTGTATGACT
TGTATACACATATGGTATCAAAGGGATAAAAGCCAATTGCTGTTACATTCCTTCACGTATTCTTTAG
CAGCACTCTGCTACTAAAGTTAATGTGTTACTCTCTTCCACATTCTCAATTAAAGGTGAGCTAACGCT
CCTCGGTGTTCTGATTAACAGTAAATCCTAAACTGTTAAATGACATTTTATTTATGTCTCCTTAA
CTATGAGACACATTTGTTACTGAATTCTTCAATATTCCAGGTGATAGATTTGTCG

FIGURE 3

GGCACGAGGCAGTCAGGGGAGGGAGCAACTGGAGCCTCAGGCCCTCAGAGTAGTCTGCCTGACCACCCCTGGAGCC
CACAGAAGCCCAGGACGTCTCCCGCGAGGCCTCCCGTGTGGCTGAGGATGGGCTGAGCAGCAGGGCCGGAGCTT
GAGGCTGAGTGCCCCGTCTGCTGGAACCCCTCAACAACACGTTCCATACCCCCAAAATGCTGGATTGCTGCCACTC
CTTCTGCGTGGAAATGTCTGGCCACCTCAGCCTGTGACTCCAGGCCGGCGCCGCTGCTGTGCCACTCTGCGCC
AGCCCACAGTGTGGCCTCAGGGCAGCCTGTCAGTGA~~CT~~CCAGGCCACGGACACTGCCATGCTCACCCCTGCTCCGCTG
GAGCCCCACCATGTCA~~T~~CCTGGAAGGCCATCAGCTGTGCCCAAGGACCAGGCCAAGAGCCGCTACTTCCTGCGCCA
GCCTCGAGTCTACAGCTGGACCTTGGCCCCCAGCCTGGGGCCAGACTGGGCCGCCCCAGACACGCCCTGCGCCA
CCGTGTCTACGCCATCCTCATCCCCAGCCACCCTTTGAGGGAGTGTTCGCAACCCCTAGTTCCGATCTT
GCCTACCTGATGGCGTCATCCTCAGTGTCACTCTGCTCATATTCTCCATCTTGGACCAAGCAGTCCTTGG
GGGTGTGGGGTGAGTGCTTCCCAGACAAGAACCAACCTTTGGTTGCTGCTGGGTATGGTACTACGGAGC
CTCATTGGTATTGTCTCCCTTGTAGTGTGTTATTTACAATCCAGGGATTGTTAGGCCATGTGTTGCTTCT
GGGAACAATTAAAAAAAAACAAAAACGAAAAGCTGAAGGACTGGGAGATGTGGAGCGACCTCCGGGTGAG
TGTGGCGTCATGGAAGGGCAGAGAAGCGGTTCTGACCACAGAGCTCCACAGCAAGTTGTC~~CC~~AAAGGGCTGCACAGT
GGTATCCAGGAACCTGACTAGCCAAATAGCAAGTTGCATTCTCACTGGAGCTGCTCAAATCAGTCATATT
TTTGAGTTGCTCTTACTATGGGTTGCTAAAAAAAAAAATTGGGAAGTGAGCTTCATTCTGTGGGTA
AATGTGTGTTGTTCTTTGAATGTCTGCCACTGGTTGCA~~G~~TAAGTGTCTGTATTCAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 4

GAAGCGCGCTCCGGGGAGGTGTTGCAGCCATGGCACGGCAGCCGGCGACCTACTTCAGCGAGGCAGTCTGTT
 CTGGTTCACAGTCATCACCCCTCAGCTTGCTACTACACATGGGTTGCTCTGCCCTCAGAGTATCCCTATCAGA
 ACCTTGGGCCCTGGGCCCTCACTCAGTACTTGGTGGACCACCATCACACCCCTTGCATGGGTATTGGCTT
 GCCTGGCTGATTCATGTGGAGAGTCCTGTATGCCATAGCATTGTGCAAGCATAAAGGCATCACAAGTGGCGGGC
 TCAGCTACTCTGGTCCTACAGACTTCTTGGATAGCGTCTCACCATTTGATTGCTTACAAACGGAAGC
 GCCAAAAACAAACTTGAAGTTGTCTGAAAGCTTGCTCTACACTTACATTCATCCTCACCCCTTTTGTGGGGT
 AGAGGAGGTGCAGTAATTACTCAGTGATCTTCTACTTCTAGAAACTGTCCTCAAAGCTTTAACGACCCCTC
 GTTAGTCAGTTCTCTTATATGCTCTGGTGAAGACAGCTTACAGTCTTGGCTGTGTTGACAGGAAA
 GATTTAGCTTCAATCCTATTGGCAGAGGACTTCAGCTACCTTCTACAGTCTTGGCTGTGTTGACCCCTCG
 TGTGCTCTGAGCTAACGCCACATACTAAACTGACTTTGGTTGTATACCCCTGCTCCGCCTCTGATGAAA
 CACTTACCCCTCACAACCACCATTTCTCTCCAAAGCTTTCCACCTTGCTGCACTAAGATAAAGTGA
 ACTTCCACTATATGTCAATTCCACACACATTATTAGGTACCTGTGAGGTAGGATCCTATCCTCTCAA
 ACTTCCATTGCTACAGAGAAAAGATAAGGAAGATGAGCAAGTGCCTGGAATGGGCAGGCTGAGCAGTCAC
 CAGGCTGAGAACCTGGAGGGGAGACTGCAGAGTGCCTCCCTGATGCTGCAGCCGAAGTGATCCTCC
 CCTACCTGGCCACTGTGCTCTGCAGTGTGCAGGGCCTGATGGCACTGCTAGATTGCTCCTCAGCTCAG
 GCTTAAACAGCTTACCTTCCCCTCAGCACCTGCCCCACTATCTGCACACAGGTGCTCTAACCATGTTATT
 GAACAAAGGAGGGAAACTGATTCACTTCACTTGTCATTATCATTCAAATTGTTATGTGAAA
 ATGGCACAACCCATTGGGTACCCCTCACCCAAAATAAAAGCCAAGTCTACCTTGACTGGTACCA
 CCTTTGTGGTTCTGGTGTGAGAACCTTATCTTCTACCTTCTATTCTCAATCACTCTCC
 AAAAGTGTGCTTTCCAGCTCTGATTTACCTCTCCTGTTCACTTAGTT
 ATACTTATTATTGCTCACAGGCTGGGAGGCAGAATGACTCTGTCACCACTAGGAGCCATTAGGG
 CCTCTCCCTGAGGACTGCCTGCTTCTGGGACACTAGCCCTCATTCCCTCTGTGGTACAGTGG
 GCAATTATTGTAT
 TAAGCAAACATTATGGAAACAACCGCTCCGAAAACGGAGCCCCAAGTAAAGCACA
 ACCCTGAAAGATTATGA
 ACTATGAATTGTCCTAGTAGAGATAAATTCTGCAAACATATCTCAGTCTCC
 CTGTTCTGTTGATTAAGA
 AGTTCCCTTTGGTAAGGAAAAGGATTAAACCATAGAGTTAGGCATCATGG
 AAATTCAAACCA
 GAGATTCTTAATA
 CCTGGTCTCCTCAAAGAGAAATAAACAGTAATAGGGTGTGG
 GACAATATGGCAGATTATTGAATGAAATTG
 ATTAACCTGAATAAAATGCTGTGAATTTC

FIGURE 5

GGCACGAGGCCGAGCGGACTGCCCTTCCCAAGATGGCGTCGAAGATAAGGTTCGAGACGGTGGATGTTGCAGCTGA
 TCATGCAGTTGGGTCGGTGCTGCTCACACGCTGCCCTTGGGCTGCTTCAGCCAGCTCATGCTGTACGCTGAG
 AGGGCTGAGGCACGCCGAAGCCCACATCCCAGTGCCTACCTGTATTCGACATGGGGCAGCCGTGCTGTGCGC
 TAGTTCATGTCCTTGGCGTAAGCGCGCTGGTCGCGCTGGGGCCGACTCCAATTGCCATTAGCACCTACG
 CCGCCTACATCGGGGCTACGTCCACTACGGGACTGGCTGAAGGTCCGTATGTACTCGCGCACAGTGCACATC
 GCGGCTTCTTGTGTTGGCCAGCGGTGCTGGGAGCTGTACGCCGAAACCTCGCAGCCGCTCCCTGCAGTCAC
 CGGCCAGGTGTTCTGGTATCTACCTCATCTGTGTCGCCTACTCAGCAGCAAGGAGGACCGGCTGGCGT
 ATCTGAACCACATCTCCAGGAGGGAGCTGATGATCCAGCTGTTCTCGTGTATGGCATCCTGCCCTGGCCTT
 CTGTCAGGCTACTACGTGACCCCTCGCTGCCAGATCCTGGCTGTACTGCTGCCCTGTCATGCTGCTCATTGATGG
 CAATGTTGCTTACTGGCACAAACACGCGCGTGTGAGTTGGAACCAGATGAAGCTCCTGGAGAGAGTGTGGCA
 TCTCGGAACTGCTGTCATCCTGCCACTGATGGCTGATTTATGGCAAGAGGCTGAGATGGCACAGGGAGCCAC
 TGAGGGTCACCTGCCTCCCTGCTGGCCAGCTGCTTTATTATGCTTTGGTCTGTTGATCTT
 TGCTTTTAAAATTGTTTGCAGTTAACAGGGCAGCTCATTGTCAAATTCTGGCTCAGCGCTGGAGGGC
 AGGAGCCCTGGCACTAATGCTGTACAGGTTTCTGTTAGGAGAGCTGAGGCCAGCTGCCACTGAGTCCTG
 TCCCTGAGAAGGGAGTGGCAGGGCTGGGATGCGGCTACTGAGAGTGGGAGAGTGGGAGACAGAGGAAGGAAGATG
 GAGATTGAAAGTGGCAAATGTGAAAAATTCTCTTGAACCTGGCAGATGCAGCTAGGCTCTGCAGTGCTGTTGG
 AGACTGTGAGAGGGAGTGTGTGTTGACACATGTGGATCAGGCCAGGAAGGGCACAGGGCTGAGCACTACAGAA
 GTCACATGGTTCTCAGGGTATGCCAGGGCAGAAAACAGTACCGGCTCTGTCACTCACCTTGAGAGTAGAGCAGA
 CCCTGTTCTGCTGGCTGTGAAGGGTGGAGCAGGCAGTGGCCAGCTTGCCCTGCTGTCTGTTCTG
 CTCCATGGTGGCCTGGTGGGGGTGGAGTCCCTCCAAACACCAACAGTCCCTAAAAAATAAACATTAA
 TATAGACAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6

GGCCGCGGCTGCCCTTGGCCCTCTTATCAGGATGAAAACGTTCTGTTGGTGTCTGGGCCCTGCTGGCCTTGAT
CCTTGCCCAGGGTCCCGGAAGAGTTGAGGTTCTATTGGCCAAGTCAGGCCCTGGTGGAGTTGGACAGT
CCCTAGTGTGCAACTGCAGCACTACTTGCCCAGACCCAGGACCCAGTGAATTGAGACCTTCTAAAGAAAACCTCAG
GTGGACAAAGGGCCTCAGTGGAAAGAGTTCTCTGGAGGATGTCACAGAGAATTCCATCCTGCAGTGCTTCTTC
TTGTGCAGGGATTCAAAGGACACAAGCCTTGGCATCACTGTGTATCAGCCACCAGAGCAAGTGATCCTGGAGCTGC
AGCCTGCCCTGGGTGGCGTGGACGAAGCCTCACAGTGAAGTGTATGTACCCAGTGTAGCACCCCTGGAGAGTCTC
ACCCTGCCCTCTCCAGGGTAACCAAGAACTGCATAGAAAGAACTTACGAGCTGGCTGGCCTCCAAAGAGC
TGAAGTCATCATCAGTGTAGGCCAAAAGGAGAACATGACAGATGCAATTCTCCTGCCATGCAGAACTGGACTTGA
TTTGCAAGGTGGGAGGCTTTCAAGGCAGCTACCCATCAGAATAGTCGGATCTTGAAATTCTCAGAGTCCC
CACATCTGGTCTCTCCCTTTGGAGGCTGGATGGCGGAGACTGTGAGCTGCGAGGTGGTAGGGTGGTCTCCAGC
CAAAGAAGTTATGTTCCACATGTTCTGGAAAGACCAAGAGCTGAGCTCCCTCCCTGGAGGGGACACAGCAT
GGGCAATGCTACCATTGGACCATGGAGGCTGGTGTACAGGAACCTGTCTTGCTTGCATCTGGTCAATGGAA
CAGAAGACAAGAAAGCTAGTGCATAGCTACAGCTCCCTCCACCAATCCTGGAGCTAAAGAACATACCCATTGGC
AGGGACCGACATTAATGTGACCTGCTCAGGCATGTATTAACATCACCCAGCCACTCTCGGCTCAGGGAGCCC
CAGACCTCCCTGCTGGGAGCCTGCCCTACTACTGCCAGGGAGGAAGATGATGGCXGAAATTCTCCTGC
GAGGCCTTTGGTGTGCAGGGTCAGCGTTGATGAAAACCACTGTGATCCAGCTCCATATCCTAAAGCCACAGTT
AGAGGAATCCAGTGCCTGGCAAACAGACCTGGCTGGAAGGGATGGAACACACGGCTGCCCTGCGTCCAAAGGGAA
ACCCAGCTCCAGCCTGGTGTACCTGGAATGGGTGGTCTTGACCTTGAAGTGCCACAGAAGGCAACCTAGAAC
CACACTGGAACCTACCGCTACACAGCCACTAACAGCTGGCTCTGTCAGCAAAGACATTGCTGT

FIGURE 7

CCACGCGTCCGTTCTGAGGTGCATTCTTTTGATGAGAGGCATCTAGGTACCATCCCTGACCTGGTCCT**CATG**
 CTGCCGAGGCTGTTGATCTGTGCTCCACTCTGTGAACCTGCCAGCTGTTTGATGCCAGCCCCCTCCA
 TCCCACAGAGGGGAGCCAGTGACCTGACGTGTAAGATGCCCTTCTACAGAGTCAGATGCCAGTTCCAGTTCT
 GCTTTTCAGAGACACCCGGCCTGGGCCAGGCTGGAGCAGCTCCCCAAGCTCCAGATCGCTGCCATGTGGAAA
 GAAGACACAGGGTCATACTGGTGCAGGCACAGACAATGGCTCAAAGTCTGAGGAGCAGGAGATCCCAGATAAA
 TGTGCACATCCCGGTGTCGCCAATCCTCATGCTCAGGGCTCCAGGGCCAGGCTGCAGTGGAGGATGTGCTGG
 AGCTTCACTGTGAGGCCCTGAGAGGCTCTCCTCCAATCCTGACTGGTTTATCAGGAGATATCACCCTGGGGAGC
 AGGTCGGCCCCCTCTGGAGGAGGAGCCTCCTCAACCTTCCCTGACTGAAGAACATTCTGGAAACTACTCCTGTGA
 GGCAACAATGGCTGGGGGCCAGGCCAGTGAGGCGGTGACACTCAACTCACAGTGCCTACTGGGCCAGAACGA
 ATCATCTTACCTCAGGAGTCATTGAGGGCTGCTCAGCACCCCTGGTCCAGGCCACCGTGGCCTTATTATTTGCTAC
 GGCCTCAAAAGAAAAATAGGAAGACGTTCAGCCAGGGATCCACTCAGGAGCCTCCAGCCTAACCCAAAGAGTTCAC
 CTACCTCAACTCACCTACCCAGGGCAGCTACAGCCTATATATGAAAATGTGAATGTTGTAAGTGGGATGAGGTT
 ATTCACTGGCGTACTATAACCAGCCGGAGCAGGAATCAGTAGCAGCAGAAACCCCTGGGACACATATGGAGGACAAG
 GTTCCCTAGACATCTATTCCAGGCTGAGGAAAGCAAACATTACAGATGTGGACTATGAAGATGCTAT**GTAAAGTTA**
 TGGAAGATTCTGCTCTTGAAAACCATCCATGCCAACGCTCAGGCCAGGCTGATATGTTCTTCAGAGATCCTGGGCA
 TTAGCTTCCAGTATACTCTGGATGCCATTCTCCATGGCACTATTCCATCTACTGTGAAGTGAAGTGGC
 GCAGCCCTGAAGAAACTACCTAGGAGAACTAATAGACACAGGAGTGACAGGGACTTGTATCAGAACCGAGATTCT
 GCCGGCTCCTTGAAAACAGGTATATTGTCCTCTGTTACAAGAGGAAACAAGATGGAATAAAAGAAATTGGG
 ATCTTGGGTTGGAGGGACAGTGAAAGCTAGAGCACATGAACACTCAAGGTTAGTGAUTCTGCAGGACTTCACAGAGAGA
 GCTGTGCCCATCATTCAAGTCAAGTGCTTCTGCCAGACAGCACAGAACTCCAGCCCCGCTACTTACATGGATC
 ATCGAGTTCCACCTAAAATATGATTCTATTGAGTCAGTGTACCTGTTACCAATTAGAACACTAAAACAAAGTTACAT
 AAAAGTTATTGTGACTCCACTTAATTAGTGAAGCTTATAAATACAGTGTGCTCTCTTATTCAACAAATTGGCAACCTATACCACATCCAAAT
 TATGTATCTATTACAGCCCTAGAACGCTTATAAATACAGTGTGCTCTCTTATTCAACAAATTGGCAACCTATACCACATCCAAAT
 GTAATATGGTTGAAAACCTGTATCTTAATTATTTTTAAATTGAGACAGGGCTCAGTGTCAACTCTGCAACTCAATCTG
 GAATGCAGTGGCACAATCTGCCTCACTGCAACGCCCTGCCTCTCAGGCTCAAGCAAACCTCTCACCTCAGCCTGCTG
 AGTAGCTGGACTACAGGCACATGCCACAAACTGGCATTGGCTTACGTAGAGAACAGATTCAACCGTT
 GCCCAGGCTGGCTCAAACCTCTGGCTCAAGCAATGTATTGAATTAAACCAACAGGCACTCACTCTTATGAAT
 TAATAAACATTGGAGGTATATAAGTAAAAGTTAAAGTCTTCCTGTAAGTAAACACAAATGTTAACTATTGTTA
 AAAACTTACAGGTAGCTCTAGATATTGTTATGTACTTATGCATACATGTAAGTATATAAAC
 TTTAGAAGTGTACCTATCTAACAAACTATTGAAATACTTCAAACTGTAAATAGATCTATTACTATTGTTA
 AGTCTCTAGTAGTGTGTTATAGATAATCATAACTTTCTTTTATTGTTAGTAAATATGCACAAACATAA
 AATTGATCATTAAACCATTAAAGTGTACAATTCACTGTCAGTGGCATTAAAGTACTATCATAATATATTAAATCCTCTC
 ATCACTGGTGGACATTAAGGAGACTCTCAAAAAATTCAATTATAAAAACAAAGTCAAAACAAATGTCTTGTACTA
 GCATATTATGGCACTCCTGCTGGATTATCTGAAGGATAAATTGTTAAATCTAGTATTGCTAGATTGCAATTAAA
 TATTCTGTTAAATAGTCAAAAAAAAAAAAAA

FIGURE 8

CTCAATCAGTTATGCAGAGAAGAAGCTACTGAGCTACTGCTGGTGTAGGCAAGTGCTGCTTGGCAA
TCTGGGCTGACCTGGCTCTCCAGAACCTCTCCAGAACCTGGAGCAGGCTTCAGCTGCTGCTGGCGTCC
TTGCTGCCCTTGCTCCAGTCTGTGGACAATCTGCAGCTGACACAAACCTGTGATTTCGTCATCCTCATGGAC
CACATTCTCAAAGGAGAGAGAGTGACTCTGAATTGCAATGGATTCTAGTTCTATGCAACAGAGAAAACAACATGGT
ATCATCGGCACTACTGGGGAGAAAAGTTGACCCCTGACCCAGGAAACACCCTCGAGGTCGGGAATCTGGACTGTAC
AGATGCCAGGCCGGGCTCCCACGAAGTAACCCCTGTGCGCTTGCTTTCTCAGACTCCTAATCCTGCAGGC
ACCATATTCTGTGTTGAAGGTGACACATTGTTCTGAGATGCCACAGAAGAAGGAAAGAGAAATTGACTGCTGTGA
AATATACTTGGAAATGAAACATTCTTCCATTCTAATAAAAGCTGGATCTTCTATCCCACAAGCAAGTCAAAT
AACAATGGCAATTATCGATGCATTGGATATGGAGATGAGAATGATGTATTAGATCAAATTCAAATAATTAAAAT
TCAAGAAGACTATTCCACATCAGAGCTGAAAGCTACAGACTCTCAGCCTACAGAGGGGAATTCTGTAACCTGAGCT
GTGAAACACAGCTCCTCCAGAGCGGTAGACACCCCCTCACTTCACCTCTCAGAGATGGCGAGGTCATCCTG
TCAGACTGGAGCAGTACCCCGAACTCCAGCTCCAACCCGCTGGAGAGAAAACCTCAGGATCTATTGGTGTGG
TGAAAACAGTGGAGGGTAACATCCACAGCAGCTCCCTCGTACAGATCCATGTGCGAGGGATCCTGTCTGGGG
TGCTCCTGGAGACCCAGCCCTCAGGGGGCAGGGCTGTGAAGGGGAGATGCTGGCCTTGTGCTGCCGTGGCTGAA
GGCACAGGGGATACACATTCTCCTGGCACCGAGAGGACATGCAGGAGAGTCTGGGGAGGAAAACCTCAGCGTCCCT
GAGAGCAGAGCTGGAGCTCCCTGCCATCAGACAGAGCCATGCAGGGGAGATACTACTGTACAGCAGACAACAGCTACG
GCCCTGTCCAGAGCATGGTGTGAATGTCAGTGTAGAGAGAGACCCAGGCAACAGAGATGGCCTTGTGCCCGGGGA
GCCACTGGAGGGCTGCTCAGTGTCTTCTCCTGGCTGTGGCCCTGTTCACTGCTGGCCTGGAGGAAGTCAGG
AGTTGGTTCTTGGAGACGAAACCAGGCTCCCTCCAGGCCAGGAGAGTCTCCATTCCATCTGCCCTG
CCCAGGTGGAGCTCAGTGTGATGTTGATGTCACCCCCAAAAGGGAGATTGGTATACTCTGAGATCCAGACT
ACTCAGCTGGGAGAAGAAGAGGAAGCTAATACCTCCAGGACACTCTAGAGGATAAGGATGTCAGTTGTACTC
TGAGGTAAGACACAACACCCAGATAACTCAGCTGGAAAGAGTCAGCTCTAAGGATGAAGAAAGTTAAAGAGAATGAAA
AGTTACGGGAACGCTCTACTCATGTGATTCTCCCTGTCAAAGTCCCAGGCCAGTGCAGTCCTGCGGCACCTG
GAATGATCAACTCATTCCAGCTTCTAATTCTCATGCATATGCATTCACTCCAGGAATACTCATTGTC
CTGATGTTGGGATGGAATGCCCTGAAAGACTTCACTAAATGACCAGGATCCACAGTTAAGAGAAGACCCCTGTAG
TATTGCTGTGGGCTGACCTAATGCATTCCCTAGGTCTGTTAGAGAAGGGGATAAAAGAGAGAGAAGGACTGT
TATGAAAAACAGAAGCACAATTGGTGAATTGGGATTGCAAGAGATGAAAAGACTGGGTGACCTGGATCTG
TTAACATCTACAACCATTGTCACTGGAGACTCACTGCATCAGTTGTTAACTGTGAGTGGCTGCACAGGCA
CTGTGCAAACAATGAAAAGCCCTTCACTCTGCCTGCACAGCTACACTGTCAGGATTGCAAGATTAAAGAA
CCCATCTGGAATGTTACAGAGAGGAAATTAAAAGAGGACATCAGAAGAGCTGGAGATGCAAGCTCTAGGCTGC
GCTTCCAAAAGCAAATGATAATTGTTAATGTCATTAGTCACAAAGAGATTGCAACATTAGAGAAAAGAGACACAAA
TATAAAATTAAAAGCTTAAGTACCAACTCTCCAAAACTAAATTGAACTTAAATTAGTATAAAACTCATAATAAA
CTCTGCCCTTAAAAAAAGATAATTCTCCTACGTCCTGTCACTGAAATAATTACCAACCCCTTAGCAATAAGCACT
CCTGCAAGAGGAGTTTATTCTCAAATTACCATCCCTCAGGAAATTAGGTTGCTTCTGTTAGGAAGCTGT
GTCTTGTAGTTACTAATTAGTTATGAGAATAATTGTCACAAATGAGAAGAGGAAATAAAAGAAAATAGGAAGCC
ACAAATTGTTATGATTTCATGATACACCTACTGGTAAATAATTGACAAAAGCAGCAGCCAAATATTAGAGGT
CTCCTGATGGAAGTGTACAATACCACTACAAATTATCCATGCCCAAGTGTAAACTGAACTCATTCAAGTCTT
CTAAGTAATACTGTTTATGAAAATGAGGAAAGGAAATTGTTAAATAACATTATGGGATGCAACCCAG
AAAACATAAAACTGAGAAAAGTTCTATAGGGCAAATCACCTGGCTCTATAACAAATAATTGGGAAAAAAATGAAA
TAAAAGAAGAGAGGGAGGAGAAAGGGAGAGAGAAGAAAAGAAAATGAGAAGAAAAGTAATTAGAATATTTCACAA
TAAAGAAAAGACGAATATTAGGTGACAGATATCCAACTACGCTGATTGATTTACAAATTATGAGTGTAT
GAATTGTCACATGATCACCCCAAAAAAGAGAAAAGAAAATAGAAGACATATAAATTAAATGAGACGAGACA
TGTCGACCAAAAGGAATGTGTTGGTCTGTTGGATCCTGACTCAAATTAGAAAAAAATAACTACCTACGAAATA
CTAAGAAAATTGTACTAATATTAGAAATTGTTGTGTTGGATATAAGTGTAGTTATTGAGTGT
TTTATAAAAGCAAAAGGATATTCACTTCAGCGCTTACTGAAGTATTAGATTAAAGCTTATTAACGTA

FIGURE 9

GCCGAGCTGAGCGGATCCTCACATGACTGTGATCCGATTCTTCAGCGGTTCTGCAACCAAGCGGGTCTTACCCC
 CGGTCCCTCCCGTCTCCAGTCCTCGCACCTGGAACCCCAACGTCCCCGAGAGTCCCCGAATCCCCGCTCCAGGCTA
 CCTAAGAGGATGAGCGGTGCTCGACGGCGGGGAGCCCTGATGCTCTGCGCCGCCACGCCGTGCTACTGAGCGC
 TCAGGGCGGACCGTGAGTCCAAGTCGCCGCCTTGCCTGGAGAGATGAATGTCTGGCGACGGACTCC
 TGCAGCTCGGCCAGGGCTGCGGAACACGCGGAGCGCACCCGAGTCAGCTAGCTGAGCGCGCTGGAGCGGCCCTGAGC
 GCGTGGGGTCCGCCGTCAAGGAACCGAGGGTCCACCGACCTCCCGTAGCCCTGAGAGGCCGGTGGACCCCTGA
 GGTCTTCACAGCCTGCAGACACAACCTCAAGGCTCAGAACAGCAGGATCCAGCAACTCTTCCACAAGGTGGCCAGC
 AGCAGCGGCACCTGGAGAACGAGCACCTGCGAATTCAAGGCTCAGCATCTGCAAAGCCAGTTGGCCTGGACCACAAGCAC
 CTAGACCATGAGGTGGCAAGCCTGCCGAAGAAAGAGGCTGCCAGATGGCCAGTTGACCCGGCTCACAA
 TGTCAGCCGCCTGCACCGGCTGCCAGGGATTGCCAGGAGCTTCCAGGTTGGGAGAGGCAGAGTGGACTATTG
 AAATCCAGCCTCAGGGTCTCCGCATTTGGTAAGTGAAGATGACCTCAGATGGAGGCTGGACAGTAATTCA
 AGGCACGATGGCTCAGTGGACTTCACCGGCCCTGGAAAGCCTACAAGGCGGGTTGGGATCCCCACGGCGA
 GTTCTGGCTGGGCTGGAGAACGGTGCATAGCATACGGGGGACCGAACAGCCGCCTGGCGTGCAGCTGCGGACT
 GGGATGGCAACGCCAGTTGCTGCAGTTCTCGTGCACCTGGTGGCAGGACACGGCTATAGCCTGCAGCTCACT
 GCACCCGTGGCGGCCAGCTGGCGCCACCACCGTCCCACCCAGCGGCCCTCCGTACCCCTCTCCACTGGGACCA
 GGATCACGACCTCCGCAGGGACAAGAAACTGCGCCAAGAGCCTCTGGAGGCTGGTTGGCACCTGCAGCCATT
 CCAACCTCAACGCCAGTACTCCGCTCCATCCCACAGCAGCGGCAGAACGCTTAAGAAGGAAATCTCTGGAAGACC
 TGGCGGGCCGCTACTACCCGCTGCAGGCCACCACATGTTGATCCAGCCCAGGCAGAGGCAGCCCTCCTAGCG
 TCCTGGCTGGCCTGGTCCCAGGCCACGAAAGACGGTGAATCTGGCTCTGCCAGGGATGTGGCCGTTCCCTGCC
 TGGCAGGGGCTCCAAGGAGGGCATCTGGAAACTTGTGGACAGAGAACGACCACGACTGGAGAACGCCCTTT
 CTGAGTGCAGGGGGCTGCATGCGTTGCCCTGAGATCGAGGCTGCAGGATATGCTCAGACTCTAGAGGCCTGGAC
 CAAGGGCATGGAGCTTCACTCCTGCTGGCAGGGAGTTGGGACTCAGAGGGACCACTGGGGCCAGCCAGACTG
 GCCTCAATGGCGACTCAGTCACATTGACTGACGGGACCAAGGGCTTGTGGTGGCGAGAGCGCCCTCATGGTCTG
 GTGCTGTTGTGTAGGTCCCTGGGACACAAGCAGGCCAATGGTATCTGGCAGGAGCTCACAGAGTTCTTGG
 ATAAAAGCAACCTCAGAACAC

FIGURE 10

GATGTGCTCCTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTGGAATTGAGGAAACT
TCTCTTGATCTCAGCCCTGGTGGTCCAGGTCTTCATGCTGCTGTGGGTGATATTACTGGTCCTGGCTCCTGTCA
GTGGACAGTTGCAAGGACACCCAGGCCATTATTTCCCTCAGCCTCCATGGACCACAGTCTCCAAGGAGAGAGA
GTGACCCCTCACTTGCAGGGATTCGCTTCACTCACCAAGAAAACAAATGGTACCATCGGTACCTGGAAAGA
AATACTAAGAGAAACCCCAGACAATATCCTTGAGGTTAGGAATCTGGAGAGTACAGATGCCAGGCCAGGGCTCCC
CTCTCAGTAGCCCTGTGCACTTGGATTTCTCAGAGATGGGATTTCTCATGCTGCCAGGCTAATGTTGAAC
CTGGGCTCAAGTGATCTGCTCACCTAGGCCTCTCAAAGCGCTGGATTACAGCTCGCTGATCCTGCAAGCTCCACT
TTCTGTGTTGAAGGAGACTCTGTGGTCTGAGGTGCCGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACA
AGAATGATAATGTCCCTGGCATTCTTAATAAAAGAACTGACTCCAAAAAAAAAAAAAA

FIGURE 11A

CGCCCGCGCGCTGCAGCCCCATCTCCTAGCGGCAGGCCAGGGCGGGAGGGAGCGAGTCGCCCGGAGGTAGGTCCAG
 GACGGGGCGCACAGCAGCGAGGCTGGCCGGGAGAGGGAGGAAGAGGAATGGCAGGGCACGCCCGAGCCATGGG
 CCAGGCTGCTCTGGCAGCCTTGATCAGCGTCAGCCTCTCTGGGACCTGGCAAAACCGCTGCAAGAAGGCCCAGTG
 AAGAGCTGCACGGAGTGTGTCGTGATAAGGACTGCGCTACTGCACAGACAGAGATGTTAGGGACCGGGCGCTG
 CAACACCCAGGCGGAGCTGCTGGCCGGGCTGCCAGCGGGAGAGCATCGTGGTCATGGAGAGCAGCTTCAAATCA
 CAGAGGAGACCCAGATTGACACCACCCCTGCGGCGCAGCCAGATGTCCCCCAAGGCCTGCGGGTCCGTCGCGGCC
 GGTGAGGAGCGGCATTTGAGCTGGAGGTGTTGAGCCACTGGAGAGCCCCGTGGACCTGTACATCCTCATGGACTT
 CTCCAACCTCATGTCCGATGATCTGGACAACCTCAAGAAGATGGGCAGAACCTGGCTCGGGCTGTAGCCAGCTCA
 CCAGCGACTACACTATTGGATTGGCAAGTTGTGGACAAAGTCAGCGTCCCGCAGACGGACATGAGGCCTGAGAAG
 CTGAAGGAGCCCTGGCCAACAGTGACCCCCCTCTCCTCAAGAACGTATCAGCCTGACAGAACAGATGTGGATGA
 GTTCCGGAATAAAACTGCAGGGAGAGCGGATCTCAGGCAACCTGGATGCTCTGAGGGCGGCTTCGATGCCATCCTGC
 AGACAGCTGTGTCGACGAGGACATTGGCTGGCCGGACAGCACCCACCTGCTGGTCTTCTCCACCGAGTCAGCC
 TTCCACTATGAGGCTGATGGGCCAACAGTGTGGCATCATGAGCCGAAACGATGAACGGTGCACCTGGCACCTGGACAC
 CACGGCACCTACACCCAGTACAGGACACAGGACTACCCGCGGCGGAGGAGCTGGTGCAGCAGGGCTTCGTCGCGGCAAG
 ACATCATCCCCATTTGCTGTCACCAACTCTCTATAGCTACTACAGAGAACGCTTCAACACCTATTCCCTGTCTCC
 TCACTGGGGTGTGTCAGGAGGACTGTCACACATCGTGGAGCTGCTGGAGGAGGCCTTCATCGGATCCGCTCCAA
 CCTGGACATCCGGCCCTAGACAGCCCCGAGGCCTTCGGACAGAGGTACCTCAAGATGTTCCAGAACAGAGGAGGA
 CTGGGTCTTTCACATCCGGGGGGAGTGGTATATACCAGGTGAGCTGCGGGCCCTTGAGCACGTGGATGGG
 ACGCACGTGTGCCAGCTGCCGGAGGACAGAACGGCAACATCCATCTGAAACCTCCAGTCTCCGACGGCCTCAAGAT
 GGACGCGGGCATCTGTGATGTGTCACCTGGAGCTGCAAAAAGAGGTGCGGTGAGCTGCTGAGCTTCAGCTCAAC
 GAGACTTCGTGTGCGGACAGTGTGTGTCAGCGAGGGCTGGAGTGGCCAGACCTGCAACTGCTCCACCGGCTCTG
 AGTGACATTCCGGCTGCTGCCGGAGGGCGAGGACAAGCCGTGCTCCGGCGTGGGAGTGCCAGTGGGGACT
 TGTGTGCTACGGCGAAGGCCCTACGAGGGTCAGTTCTGCGAGTATGACAACCTCCAGTGTCCCCGACTTCCGGT
 TCCTCTGCAATGACCGAGGACGCTGCTCCATGGGCCAGTGTGTGAGCCTGTTGGACAGGCCAAGCTGTGAC
 TGTCCCCCTCAGCAATGCCACCTGCATCGACAGCAATGGGGCATCTGTAATGGACGTGGCACTGTGAGTGTGGCG
 CTGCCACTGCCACCGCAGTCGCTCTACACGGACACCATCGAGATCAACTACTCGGCGATCCACCCGGGCTCT
 GCGAGGACCTACGCTCTGCGTGCAGTGCCAGGGTGGGGCACCAGGAGAACAGAACGGGCGCAGTGTGAGGAATGC
 AACATTCAAGGTCAAGATGGTGGACGAGCTTAAGAGAGAGCCGAGGAGGTGGTGGCGCTGCTCTCCGGGACGAGGA
 TGACGACTGCACCTACAGCTACACCATGGAAGGTGACGGGCCCTGGGGCCAAACAGCACTGTCCTGGTGCACAAGA
 AGAAGGACTGCCCTCCGGCTCCTCTGGTGGCTCATCCCCCTGCTCCTCCCTGCCCTGGGGCTCTGGCCCTGCTA
 CTGCTGCTATGCTGGAAAGTACTGTGCTGCAAGGCCCTGCTGCCACTTCTCCCTGCTGCAACCGAGGTACAT
 GGTGGGTTTAAGGAAGACCAACTACATGCTGCCGGAGAACCTGATGGCCTCTGACCAACTTGGACACGCCATGCTGC
 GCAGCGGGAACCTCAAGGGCGTGACGTGGTCCCTGGAAAGGTACCAACAAACATGCAAGCGGCTGGCTTGCAC
 CATGCCGCCAGCATCAACCCACAGAGCTGGTCCCTACGGGCTGTCTGCGCTGGGCCCTTGCACCGAGAA
 CCTGCTGAAGCCTGACACTCGGGAGTGCGCCAGCTGCCAGGGAGTGGAGGAGAACCTGAACGAGGTCTACAGGC
 AGATCTCCGGTGTACACAAGCTCAGCAGACCAAGTTCCGGCAGCAGGCCAATGCCGGAAAAAGCAAGACCACACC
 ATTGTGGACACAGTGTGATGGGCCCGCTGGCAAGCCGGCTACTACACCCACTGCAGACCAAGGAGGCCGGGATGGTGG
 AGGGCCCTTCCAGCACCTCAAGGTGGCCCCCGGCTACTACACCCACTGCAGACCAAGGAGGCCGGGATGGTGG
 AGTTCAGGAGGGCGTGGAGCTGGTGGACGTACGGGTGCCCTTTATCCGGCTGAGGATGACGACGAGAACAG
 CTGCTGGTGGAGGCCATCGACGTGCCCGCAGGCAGTGCACCCCTCGGCCGCCCTGGTAAACATCACCACATCAA
 GGAGCAAGCCAGAGACGTGGTGTCTTGAGCAGCCTGAGTTCTCGGTGAGCCGGGGACAGGTGGCCCGCATCC
 CTGTCATCCGGCGTGTCTGGACGGCGGAAGTCCAGGTCTCTACCGCAGCACAGGATGGCACCGCGCAGGGCAAC
 CGGGACTACATCCCCGTGGAGGGTGAGCTGTCTCCAGGCTGGGAGGCTGGAAAGAGCTGCAAGGTGAAGCTCCT
 GGAGCTGCAAGAAGTGTGACTCCCTCTGCGGGGCCAGGTCCGGCTTCCACGTCAGCTCAGCAACCCCTAAAGT
 TTGGGGCCCACCTGGGCCAGCCCCACTCCACCATCATCAGGGACCCAGATGAACCTGGACCGGAGCTTCACG
 AGTCAGATGTTGTCATCACAGCCACCCCTCACGGCGACCTGGCGCCCGAGAACCCCAATGCTAAGGCCGCTGG
 GTCCAGGAAGATCATTCAACTGGCTGCCCTCTGGCAAGCCAAAGGGTACAGGGTAAAGTACTGGATTCAAG
 GTGACTCCGAATCGAAGCCACCTGCTGACAGCAAGGTGCCCTCAGTGGAGCTACCAACCTGTACCCGTATTGC
 GACTATGAGATGAAGGTGTGCGCTACGGGGCTCAGGGCGAGGGACCTACAGCTCCCTGGTCTGCGCACCCA
 CCAGGAAGTGCCAGCGAGCCAGGGCGTCTGGCCTTCAATGTCGTCCTCCACGGTACCCAGCTGAGCTGGCTG
 AGCCGGCTGAGACCAACGGTACAGCCTACGAGGTCTGCTATGGCTGGTCAACGATGACAAACCGACCTATT
 GGGCCATGAAGAAAGTGTGGTGAACAACCTAAGAACCGATGCTGCTTATTGAGAACCTCGGGAG

FIGURE 11B

TCCCAGCCCTACCGCTACACGGTGAAGGGCGCAACGGGGCCGGCTGGGGGCCTGAGCAGGGAGGCCATCATCAACCT
 GCCCACCCAGCCAAAGAGGCCATGTCCATCCCCATCATCCCTGACATCCCTATCGTGGACGCCAGAGCAGGGAGG
 ACTACGACAGCTCCTTATGTACAGCGATGACGTTCTACGCTCTCCATCGGGCAGCCAGAGGCCAGCGTCTCCGAT
 GACACTGAGCACCTGGTGAATGGCCGGATGGACTTGCCTTCCCAGCACCACACTCCCTGCACAGGATGACCA
 GACCACTGCTGCTGCCTATGGCACCCACCTGAGGCCACACGTGCCACCGCGTGCTAAGCACATCCTCCACCCCTCA
 CACGGGACTACAACACTCACTGACCCGCTCAGAACACTCACACTCGACCACACTGCCAGGGACTACTCCACCCCTCACC
 TCCGTCTCCTCCACGACTCTCGCCTGACTGCTGGTGTGCCGACACGCCACCCGCTGGTGTCTGCCCTGGG
 GCCCACATCTCTCAGAGTGAGCTGGCAGGAGGCCGGTGCAGCGCTGCAGGGTACAGTGTGGAGTACCCAGC
 TGCTGAACGGCGGTGAGCTGCATCGGCTCAACATCCCCAACCTGCCAGACCTCGGTGGTGTGAAAGACCTCTG
 CCCAACCAACTCCTACGTGTTCCGCGTGCAGGGCCAGAGCCAGGAAGGCTGGGCCAGAGCGCTGAGGGTGT
 CATCAC
 CATTGAATCCCAGGTGACCCCGAGAGCCCACGTGTCCTGCCAGGGCTTCAACTTGAGCAGCTGGGAGC
 CGCCACGGCCATGGTGTTCACTGCCCTGAGCCCAGACTCGCTGCAGCTGAGCTGGGAGC
 CGCCACGGGCAAT
 GGGGATATCGTGGCTACCTGGTGACCTGTGAGATGGCCAAGGAGGGCCAGCCACCCATTCCGGTGGATGG
 AGACAGCCCCGAGAGCCGGCTGACCGTGCCGGCTCAGCGAGAACGTCCTACAAGTTCAAGGTGCAGGCCAGGA
 CCACGTAGGGCTTCGGGCCAGAGCGCAGGGCATCATCACCATAAGGTCCCAGGATGGAGGCCCTCCGCAGCTG
 GGCAGCCGTGCCGGCTTCCAGCACCGCTGCAAAGCGAGTACAGCAGCATCACCACCCACACCAGGCCAC
 CGAGCCCTCTAGTGGATGGCCGACCCCTGGGGCCAGCACCTGGAGGCAGGGGCTCCCTCACCCGGCATGTGA
 CCCAGGAGTTGTGAGCCGGACACTGACCAACAGCGGAACCCATTAGCACCCACATGGACCAACAGTTCTCAA
TGACCGCACCCCTGGCCACCCCGCCATGTCCACTAGGCGTCTCCGACTCCTCTCCGGAGCCTCTCAGCTAC
 TCCATCCTGCACTGGGGCCAGCCCACCCGATGCACAGAGCAGGGCTAGGTGTCTCTGGGAGGCATGAA
 GGGGCAAGGTCCGTCTGTGGCCAAACCTATTGTAACCAAAGAGCTGGGAGCAGCACAAGGACCCAGCCT
 TGTTCTGCACCTAATAATGGTTTGCTACTG

FIGURE 12A

GCCTTCAACTACCATCCCACCCACCTGCTGAGGAGAAAAATTCTCAAGACTCAGAGCACACAGCCAGCACCAGAGGC
CCC**AT**GACCCCTGGACAGACCAGGGAGGGGCCACCATGCTGAAGACATTCACTGTTGCTCTTCGATTTCGAT
GAGTCTGGGTATGACATCGATAGTGATGGACCTCAACCGGAGTTGGGATAGAGTCCAACCTACCCCCAGGCCCTT
GGGAGAACATCACGCTTGGGCCGAAGCCCTCTCGGATATCAAGCAAGTCCCTGCTGCTGAAGGATAAGACACAA
ATGACCTGGATCCGCCCTCCCACAAGACCTCCAAGTTCTTACATCTGCTTATAGGGCCCTACTGAGTCCAATGCAGG
TCTTACCGGTGCTGACTTGGAGGAGACAGGCTGGTCAAAGGCCAGTAAAGTCTAGAGTGGAGGCACCAGGCC
AACTGCCAAGCCATCTCTGGATTCAAGGCTGAGACCCCGCTTCCCTGGGTAATGTTAACATCCTCTGCCAT
GGCTGGCTGCAGGATTGGTATTGCTGTTAAAGAGGGATATGAGAGCCTGTTGATTACCAAGTCCAACCTGG
GACAATGGCCATATTCTCATTGACAACCTGACACCTGAGGATGAAGGGTTACATCTGCCCACTCATATCCAGA
TGCTCCCCACCCCTGGTCAAGAGCCAGCAACCCCTGAAGCTGGTTAGCAGGACTCTACCCCAAACCAACTTTG
ACAGCCCATCCTGGGCCATCATGGCACCTGGAGAAAGCCTGAATCTCAGGTGCAAGGGCCAATCTATGGAATGAC
CTTGTCTAATGAGGGTTGAGAGACTTGGAGAAGTCCCTTACACAAAGAACAAAAATGAGGCAAATTCT
TCTCCAGTCTTGAGGATCCAAGATACTGGACATTACCTCTGTTTACTATGACGATCATATAGAGGTTCACTC
CTTAGTGATGCTCTGAAAATCTGGTAACTGACACTTCCCAAAGACCTGGCTACTGCTGGCCAGTGTGCTGTTG
CCAAATGGGTCAAGATGTGAGCCTACGGTGTGAGGACAGTGGATGGTGGCTTGCACTCTATAAGAAAGGAG
AAGACAAACCACTCAATTGGATGCCACCAGCATGATGACAACACATCATCTCCTCAACAATGTAACCTAC
AGTGATACTGGCATCTATAGTGCACATCTCTCACCTGGAAGACCTCATTAGGATGCATCACACAACACTGT
GGAGCTATGGTTGAGATAAGCCCCCAAACCCCTGTCAGCTGGCCAAGCACTGTGTTCAAGCTAGGAAAGG
CCATCACCCCTCAGTGGCAGTATCTCATCCAGTACTGGAATTCTCTGGAAATGGGAAGAAAGAGAAACATTCCAA
AGATTCTCAGTAAACGGAGACTTCATCATCAGTAATGTTGACGGGAAGGCACAGGGACCTACAGTGCAGCTATCG
CGTAGAGACACATCTAACATGTGGTACATCGCAGTGGCCCTGAAGCTGATGGGCCAGCAGGCTATCTCACCT
GGAATTACGTTCTGAATGAAGCTATCAGGTTGCTCTAATCATGCACTGGCCCTGCTGTTGGTAGTGTGCTGG
ATAAGGTGGAAGTGTGGAGACTCAGAATCAGAGAACGCTGGTGTGGGAACAGCTCAAGGGTCACCATGCTCTT
CATAGTCACGGCCCTCTGCTGTGGACTGTGCAATGGGTATTGATAGAAGAGACTGAAATAGTCATGCCAACCC
CTAACGCTGAGCTGTGGCAGAGACCAACTTCCCTCTGGCCCCGTGGAAGAACTAACCTCTGGTGCAGAACCC
TCTGGCTCAACTAAGGAGTTGTTGCTGAAGGATGGGACCGGGTGATGCCACTGCCCGGCTCAGAGCAGGT
CCGGGCTGCCCTCCCCCTGGGCCCTGACCCAGGCCACACCGGGAGCTACCAACTGCCATTGTCATGGAGGAGATGG
CTGATCGGAGGCCAGTGGAGGACTTGGCTGGGGACAGACATCTCCCCAAACCTGTCATTCTGCTTCC
ACAATCGGGGCCAGGAACTAACACTCCGGTGCAAGGATGGCTGCCAGGATGGGTTGCTGTATAAGGAGGG
AGAGCAAAACCTGTCAGCAACTTGGCTGTTGAAGAGAACGGCTTCTTACAATCAGAGAATGGGAGGATAAAG
ACGAAGGCAATTACAGTGCCTCACACTGAAACACTCCCTCAAGTGGCTGAGCCAGTGGCTGGTCAAGGCGCTGGAG
CTTGTCAAAAAGAAATGTACCCCTAAGGCTTCTCAAGACATGGCCAGCCCTGTGGTCAAGGCGCTGGG
GACTTTCAATTGCTCCACCCCCCACCAGCATATGAGCTTATTCTTACAAAGATGGAAGTGAATAGCATCCAGTG
ACAGGTCTGGCAAGTCCGGGGCCAGTGCAGCTCACTTCTAATCATTCGGTGGCATGGTGTGGAGGGAAAT
TACAGCTGCCGATATTGACTTTCTATCTGGCTGAGCCAGCGACCCCTGTGGAGCTGTGTTGACAGAATTCTA
CCCCAAACCACTCTCTGGCACAGCCAGGCTGTGGTGTGTTCTGGGAAGAGGTGTGATCTGCCTGCCAACGGGA
CTTCCAGGGCATGAGGTTGCCCTCTGCAAGGAGGGAGCCATGTTCCCTACAGTTGGAGTGTCTCAGGGAAC
TCAGCTGACTTCCTCTCACACTGTTGGAGCAGAGGACTCTGGAACTATAGCTGTATCTACTATGAGACAAACAT
GTCAAACAGGGGTCAATCTCAGTATGCCCTATGATCTGGGTACTGACACATCTCCTAACGCACTGGTTGTTG
CTGAGCCAGTCTGTGGTCCATGGGCAGAATGTTACTCTCTGGCTGCCAGGGCCGGTCCATGGAGTAGGATAC
ATTCTGCACAAAGAAGGAGAAGCCACTCAATGCAGCTGGGATCACCAGTAATGACGGGCATTCCCCATCAC
CAATATATCTGTTACTAGCATGGGCGTTACAGCTGCTGCTACCACCTGACTGACCCAGTCTATCAAGATAAAC
CTAGCAACACCCCTGAACTCTAGTCACAGGCTTACTCCCCAAACCCAGCCTATTAGGCCAGCCTGGTCCATGGT
GCCCTGGCAGAAATATGACTCTCAGTGTCAAGGGGAAGTGCAGACTCAACATTGTGTTGAAGGAGGGGC
TCAGGAGCCTTAGAGCAACAGAGGCCAAGTGGGACTCAGGGCTGACTTCTGGATGCCAGCAGTGGAGGAGTGAAGACT
CTGGGATCTATAGCTGTGTTATTGACTCTACTCCCTTGCAAGCTTCAAATCACAGTGTACTCCCTGGAGGACT
TGGGTGACTGATAAGCCCCCTAAACCCCTCTGTCAGGCCAGGCTTCAAATCACAGTGTACTCCCTGGAGGACT
CCTTCAGTGGCAGGGACCCCTGCCAGGTTGAATTGTTCTAGAACATGATGGAGAAGAACCTCAGCAGTTT
CAGAGGATGGAGACTTGTCAACACGTTAGAAGGAAAGCATTGAAACTACAGCTGCCAGCTACCGCCTCCAG
GCCCTACCCCTGATATCTGGTCAAGGCCAGTGTGATCCCCTGGAGCTGGTGGGGCAGCAGGGCTGTTGCTCAGGAGTG
CACTGTAGGGAACATTGTCCGAAGTAGCCTAATGTCGGTGGTTGTTGAGCCTGGGGTAGTGTAGCCATAGAGT
GGAAGAAGTGGCCTCGACTGCCAACCAGAGGCTCAGAGACAGACGGAAAGAGACCACCATGCCCTTG

FIGURE 12B

AAGAGTGTAAACCAAGAAGGAGAACCGAGGCACCCCTGCCAATTCTCCTTCATCAACCTCTCAGAGAATCTCTGTGGAA
CTGCCCGTTCCAATATAATAATCTCCTCCTTACAAGAGCTTCCTCTCTCTCTTGCTCTCAGAGACCTATAAA
TCCAACCAGTTACCCCTGCAAGTCAGCCCCATCTGCTGTTCCCTGGTCTTAATCACCTGAGCTGGTAAAGGGGATT
CTGGGAGTTGAGAGCTCTGCCAGGGTGAGATGTTCCCTGAAGAGAGGTTCCCACCCCTGTAACTCCTCACTGTACT
GATTTACTGGCGCATGAAATTCTATTAAAAATGCATTCTGAATAAAAAGAGTATTCACTATTTAACTTCATTT

FIGURE 13A

CGGGAGCGGGGGAGCGGTGGCGCGGAGAGGGCGGCTCCAGCTCGGCTCCGGCTCGGGCTCGGGCTCCGGCT
 CCGGCTCCGGCTCCGGCTCCAGCTCGGGTGGCGGGAGCGGGACAGGGTGAGGCGGCGGAGAGGAGT
 GGGAGCAGCGGCCCTAGCGGTTGCGGGGGACATCGGACCGACGGCCCTGGATAGGCAGAAAGGAGTGGAGGCC
 TGGTGCCCAGGCCCTGGTGTGAGTATCCAGCAAGAGTGACCAGGGTAAGAAGCAAAGACTCGGTTGATTGTCCTG
 GGCTGTGGCTGGCTGTGGAGCTAGAGCCCTGGAT**AG**GGCCCTGAGCCAGCCCCAGGGAGGACATGGTCCCCTGTG
 CCTGCACTGGTGTGCTGGTTGGCAGGCAGGCCATGGTACAGCAAACCTGTCTCATTAAGTCCCTGAGGA
 CCAGACTGGCTGTCAAGGAGGGTAGCCTCCTCGTGTGCCAAGCTACAGGAGAACCCAAGCCGCGCATCACATGGA
 TGAAGAAGGGAAAGAAAGTCAGCTCCAGCCTGAGGTATTGAGTTGATGATGGGGCAGGGTCAGTGTCTCGG
 ATCCAGCCATTGCGGGTGCAGCGAGATGAAGCCATCTATGAGTGTACAGCTACTAACAGCCTGGGTGAGATCAACAC
 TAGTGCCAAGCTCTCAGTGTCAAGAGGAACAGCTGCCCTGGGTCCCTCATCGACATGGGCCCTCAGCTGA
 AGGTGGTGGAGAAGGCACGCACAGCCACCATGCTATGTGCCCGCAGGCCGAAATCAGACCCCTGAGATTCTGGTT
 AAGGACTTCCTCCTGTAGACCCCTGCCACAGCACACGCCGCTACAGCAGCTGCCCTCAGGTGCTTCAGGTGCTTGAGATA
 GAGCAGTGGAGAATCCGACCAAGGGCAAGTACAGTGCTGTGGCAGCAACTCGGCCAGGCACACGTTACTCAGCCCC
 CGAACCTGTATGTGCGAGTGCAGCTGGCTGGCTTCCTCGTCCATCCCTCCAGCAGGCCAGGAGGTGATGCCAGGC
 GGCAGCGTGAACCTGACATGCGTGGCAGTGGGTGCACCCATGCCCTACGTGAAGTGGATGATGGGGCCAGGAGGCT
 CACCAAGGAGGATGAGATGCAAGTTGGCGCAACGCTGGAGCTCAGCAATGTCGTACGCTCTGCCAACTACACCT
 GTGTGGCCATCTCCTCGCTGGCATGATCGAGGCCACAGCCAGGTACAGTGAAGCTCTCAAAGCCTCCGATT
 GATCTTGTGGTACAGAGACAATGCCACCAAGTGTACCCCTCACCTGGACTCTGGAACTCGGAGCCTGTAACCTA
 CTATGGCATCCAGTACCGCGCAGCGGGCACGGAGGGCCCTTCAGGAGGTGGATGGTGTGGCCACCACCGCTACA
 GCATTGGCGGCCTCAGCCCTTCTCGGAATATGCCCTCCCGTGTGGCGTGAACAGCATCGGGCGAGGGCCGCC
 AGCAGGGCAGTGCAGGACCGCACGGAGAACAGGCCACGCCACCGCGCGTGCAGGCACGCATGTCAG
 CGCCAGCACCAGTGTGGTCACTGGAGCCTCCGAGGCCAACGCCCTGGTGCAGGGATACCGCTACTATA
 CTCCGGACTCCCAGGCCCGAACGCTGGACAAGCACAACACCGACGCCGGCTCAGGCCGTGGCAG
 CTGCTGCCCTGGCATCACCTACAGCCTGCCGTGCTGCCCTCACGCCGTGGCGATGCCCTCCAGGCCACCAT
 CCAGGTCAAGACGAGCAGGGAGTGCCTGCCAGCCAGGCCGAGCTCCAGGCCAGGTGGAGTCGGACACCAGGATCC
 AGCTCTCGTGGCTGCTGCCCTCAGGAGCGGATCATCATGTATGAACTGGTGTACTGGCGGCAGAGGACGAAGAC
 CAACAGCACAAGGTACCTCGACCCAACCTCCCTACACACTAGAGGACCTGAAGCCTGACACACTCTACCGCTT
 CCAGCTGGCTGCACGCTCGGATATGGGGTGGCGTCTCACCCACCATGGGCTCCACCGTCCGGTAAGTTGGTCCCGCCT
 CCTCCGGCCCTCCCCAGAAGGTGATGTGTGAGCATGGGCTCCACCGTCCGGTAAGTTGGTCCCGCCT
 GCGCACAGCCGCAACGGCTATCACCCAGTACTCCGTGGCCACGAGGGCGTGGACGGAGGACCCGGCG
 TGTGGTGGATGGCATCACCGTGAGCACTCCAGCTGGGACCTGGTGGGCTGGAGAAGTGGACGGAGTACCGGGTGT
 GGGTGGGGCACACACAGACGTGGCCCGGGGAGAGCAGGCCGGTGTGGCGCACCAGTGGAGCAGTGGCC
 AGCAGGGCTCCCGGAAGGTGGAGGTGGAGCCACTGAACCTCACTGCTGTGCTACTGGAGCTGCTGT
 CAGCAAGCAGCATGCCAGATCCCGCGTACCAAGGTACCTACGTGCCGTGGAGAATGGCGAGCCCCGTGGACTCC
 CCATCATCCAAGACGTATGCTAGCCAGGCCAGTGGCGCCAGAGGAGTCCGAGGACTATGAAACCAACTATCAGC
 GGCCTGACCCCGGAGACCACCTACTCCGTTACTGTTGCTGCCTATACCAACCAAGGGGATGGTGGCCAGCAAGCC
 CAAAATTGTCACTACAACAGGTGCACTCCAGGCCGCCACCATGATGATCAGCACACGCCATGAACACTGCGC
 TGCTCCAGTGGCACCCACCAAGGAACCTGCCAGCTGGCAGCTGCTGGCTACCGCTGCACTACTGCCGGCGACGGAG
 GCGCGGCCAACACCATAGATTGCCAGGATGACCAGCACTTCACAGTCACCGCCCTGCACAAGGGGACCACCTA
 CATCTCCGGCTTGCTGCCAGAACCGGGCTGGCTGGGTGAGGAGTTCAGAGAAGGAGATCAGGACCCCGAGGACC
 TGCCCAGCGGCTCCCCAAAACCTGCAATGTGACAGGACTGACCACGTCTACAGCAACTGGCTGGGACCCGCCA
 GTGCTGGCGGAGAGGAACCGCGCAGTACAGCTACCCGTGGTGTGGAGACATCAACAGCAACAGGAGCTGCA
 GAACATCACGACAGACACCCGCTTACCTACTGCCCTAAGCCAGACACCACCTACGACATCAAGGTCCGCGCAT
 GGACCAAGGAGTCTGGCCACTCAGGCCAGCATCCAGTCCCGGACCATGCCGGTGGAGCAAGTGGTGGCAAG
 AACTTCCGGTGGCGCTGCAATGAAGACGTCTGCTGCTCAGCTGGGAGGTGGCGACTCTATAAGTCAAGTGT
 GCCCTTAAGATTGTACAATGGGAGAGTGTGGAGGTGGAGGGACTCGATGCCAGGAAAGCTGATGCCAGACCTGC
 AGCCCAACACAGAGTACTGTTGTGCTGATGAACCGTGGCAGCAGCCAGGGGCTGCCAGCACCTGGTGTCCATC
 CGCACAGCCCCGACCTCTGCCCTCACAGCCGCTGCCCTGCTACATAGAGGAGGCCGCTTCGATCTCTC
 CATGCCCTATGTGCAAGACCCCTCGCTGCTCAGGTGGTTCTACATTGTTGTGGTACCCATTGACCGTGTGGCGG
 GCATGCTGACGCCAAGGTGGAGCACACCGAGGAACGGAGCTGGAGGAGCTCTAGAAGCCATGAGCAAGGCC
 GAGGAGCAGCGGGCGGGCGCAGGAGAACGTCTGAAGCCATATGTGGCTGCTCAACTGGATGTGCTCCGGACT
 GACCTTACCTGGGGACAAGAAACTACCGGGCTTCTACAACCGCCCTGTCTCGGACTTGAGCT

FIGURE 13B

ACCAGTGCTTGTGCTTGCCTCCTGAAGGAACCCATGGACCAGAAGCGCTATGCCCTCAGCCCCACTCGGATGAG
 ATCGTGGTCCAGGTGACACCAGCCCAGCAGCAGGAGGAGCCGGAGATGCTGTGGGTGACGGGTCCTGCTGGCACT
 CATCCTCATCATCCTCATTGTCATGCCATGCCATCCTCTTGTCAAAAGGAAAAGGACCACTCTCGCTCTAAGGATG
 AGCAGTCGATCGGACTGAAGGACTCCTGCTGGCCCACCTCTGACCCCTGTGGAGATGCAGGAGCTCAACTACCAAG
 ACCCCAGGTATGCGAGACCACCCACCCATCCCCATCACCGACCTGGCGACAACATCGAGGCCCTCAAAGCCAACGA
 TGGCCTCAAGTTCTCCCAGGAGTATGAGTCATGCCATCGACCTGGACAGCAGTCACGTGGGAGAATTCAAACCTGGAGG
 TGAACAAGCCAAGAACCGCTATGCGAATGTCATGCCCTACGACCACCTCGAGTCATCCTTACCTCTATCGATGGC
 GTCCCCGGGAGTGAACATCAATGCCAACTACATCGATGGCTACCGCAAGCAGAATGCCATCGGCCACGCCAGGG
 CCCCCCTGCCGAGACCATGGCGATTCTGGAGAATGGTGTGGGAAACAGCGCACGCCACTGTGGTATGACAC
 GGCTGGAGGAGAAGTCCCAGGTTAAATGTGATCAGTACTGGCCAGCCGTGGCACAGACCTGTGGCCTTATTCA
 GTGACCCCTGTTGGACACAGTGGAGCTGGCACACATACACTGTCGACCTTCAGCAGGAGTACCTCTGAGTACCCAACT
 GAAGCGTGAGCTGGCTCAGTTCATGGCCTGGCCAGACCATGGAGTGGTGTGACTGCCACGCCAGGGCTGG
 CCTTCCATCGACGGTCAAGGCCATGCCAACCCCCCTAGACGCCAGGGCCATGGTGGTCACTGCCAGGCCGGGCTGG
 CGCACCGGCTGTCATCGTGAATTGATGCCATGTTGGAGCGATGAAGCAGAGAACCGTGGACATCTATGCCA
 CGTGCACCTGCATGCGATCACAGAGGAACATACATGGTGCAGACGGAGGACAGTACGTGTTCATCCATGAGGCGCTGC
 TGGAGGCTGCCACCGTGCAGGACACAGAGGTGCCTGCCGAAACCTGTATGCCACATCCAGAAGCTGGGCAAG
 CCTCCAGGGGAGAGTGTGACCGCATGGAGCTGAGTCAGTTGCTGGCCAGCTCAAGGGCCACACGTCCCCT
 CATCAGCGCAACCTGCCCTGCAACAAGTTCAAGAACCGGCTGGTGAACATCATGCCCTACGAATTGACCCGTGT
 GTCTGCAAGCCATCCGTGGTGTGGAGGGCTCTGACTACATCAATGCCAGCTCCCTGGATGGTTAGACAGCAGAA
 GCCTACATAGCTACACAGGGGCTCTGGCAGAGAGCACCGAGGACTCTGGCGATGCTATGGGAGCACAATTCCAC
 CATCATCGTATGTCGACCAAGCTCAGGGAGATGGGAGGGAAATGCCACCAAGTACTGCCAGCAGAGCGCTCG
 CTCGCTACCAAGTACTTGTGACCGATGGCTGAGTACAACATGCCCACTGATATCCTGCTGAGTTCAAGGTC
 ACGGATGCCCGGGATGGCAGTCAGGACAATCCGGCAGTTCCAGTTACAGACTGGCAGAGCAGGGCTGCCAA
 GACAGGCAGGGATTCTGACTTCATGGCGAGGTGCATAAGACCAAGGAGCAGTTGGACAGGATGGCCTATCA
 CGGTGCACTGCAGTGCTGGCGTGGGCCACCGGGGTGTTCATCACTGAGCAGTCGCTGGAGCGCATGCGCTAT
 GAGGGCGTGGTCGACATGTTCAAGCCGTGAAGACCCCTGCCACACAGCGCTGGCATGGTGCAGACAGAGGACCA
 GTATCAGCTGTGCTACCGTGGCCCTGGAGTACCTGGCAGCTTGACCACTATGCAACGTAACCTACCGCTCC
 CTCCCTGCCACCCCGCGTGGGCTCCGGAGGGGACCCAGCTCTGAGCCATACCGCAGTCGCTGGCATGGTGC
 CTACGCAGATGTCGACTGGCAGAGCACAGCCCACGGGGATCACAGCGTTAGGACCTGGCCACACCAATCAGA
 GAGCCTAGAACATCCCTGGCAAGTGGATGGCCAGGCCAGGGACTGTGGCCCTCTGTTCTGGCACAGACCCACCTG
 GAGCCGCTTCAAGCTCTGTTGCGCTCCGCATTCTCATGCTTCTCATGGGTGGGTTGGGCAAGGCT
 CCTTTTAATACATTAAGTGGGTTAGACTGAGGGATTTCAGGCTTCTGAGTTTGTGTTTTTTGTATGACTCTGCTGA
 TCTGAGAATGGGCCACTGTAGGGGTTGGGTTATTTGTTGTTTTTTGTATGACTCTGCTGA
 AGGACAGAACATTGCCCTCCTGCAAGAGCTGGGCTGCCAGCCTGAGCGGAGGCTGGCGTGGCCGGAGGCA
 GTGCTGATCCGGCTGCCCTCAGCCCTCAGACGAGATCCTGTTCAAGCTAAATGCAGGGAAACTCAATGTTTT
 TAAGTTTGTGTTCCCTTAAAGCCTTTTTTAGGCCACATTGACAGTGGTGGGGGGAGAAGATAGGGAACACTC
 ATCCCTGGTCGTCTATCCCAGTGTGTTAACATTACAGCCCAGAACACAGATGTCCTGGGAGAGCCTGGCAA
 GGCATTCCCTCATCACCATCGTGTGAAAGGTTAAAACAAAAACACAAAAAATAAAAACAAAAAAC
 AAAAAACCCAAAAAAAGAGTCAGCCCTGGCTCTGCTCAACCCCTCAAGAGGGAAAGCAACTCCG
 TGTGCCTGGGTTCCCGAGGGAGCTGGCTGACCTGGGCCACAGAGCCTGGCTTGGTCCCCAGCATTGCA
 TGTTGTTGAGGCTGTGGGCTGGCTGTGGCAAGGTGAATAGCACAGGTTAGGGTGTGCGCACACC
 CCATGCACCTCAGGGCCAAGGGGGCGTGGCTGCCCTTCAAGGTCCAGGCCAGTGGGCTGGTAGCACATGTC
 CCTCAGAGCAGGGGCCAGATGATTTCTCCCTGGTTGCACTGTTCAAAGCCCCGATAATCGCTCTTCCA
 CTCCAAGATGCCCTCATAAACCATGTCAGAACAGACTGGACTCTATCAATGTTACTCTAATCAGTC
 CCCAGCTGCTGAGGGCAGGGAGAGCGCCTCTCCCTGGCAGCGCTATCTAGATAGGTAAGTGGGGGGGGAA
 GGGTGCATAGCTGTTTAGCTGAGGGACGTGGTGGCACGCCAGTCCCCAAACCTAGCTAGGCTAAGTCAAG
 ATCAACATTCCAGGGTTGGTAATGTTGGATGAAACATTCAATTTCACCTTGATGCTAGTGTAGAGTTCA
 ACTGCTGTTCTATTGTTAAGAAAAACTACAGCATATTGCATAATTCTGATGGTAATAAATTGAATAATC
 AGATTCT

FIGURE 14

GGAGAGGGTGCAGGGCGAATCCGAGCCGAGCGAGAGGAATCCGGCAGTAGAGAGCGGACTCCAGCCGGCGAACCTGC
 AGCCCTCGCCTGGGACAGCGCGCTGGCAGGCAGCCAAAGAGAGCATCGAGCAGCGAACCCGCGAAGCCGGCCC
 GCAGCCGCGACCCCGCGAGCCTGCCGCTCTCCGCCGCGGTCCGGCAGCATGAGGCGCGCGCTCTGGCTCTG
 GCTGTGCGCGCTGGCGCTGAGCCTGCAGCTGCCCTGCCGCAAATTGTGGCTACTAATTGCCCTGAAGATCAAG
 ATGGCTCTGGGATGACTCTGACAACCTCTCCGGCTCAGGTGCAGGTGCTTTGCAAGATATCACCTGTCACAGCAG
 ACCCCCCACTTGAAGGACACGCAGCTCCTGACGGCTATTCCCACGTCTCCAGAACCCACCGGCCTGGAGGCTAC
 AGCTGCCTCCACCTCACCCCTGCCGGCTGGAGAGGGGCCAAGGAGGGAGAGGCTGTAGTCCTGCCAGAAGTGGAGC
 CTGGCCTCACCGCCGGAGCAGGAGGCCACCCCCCGACCCAGGGAGACCACAGCTCCGACCACATCAGGCC
 TCAACGACCACAGCCACCACGGCCCAGGAGGCCACCTCCCACCCCCACAGGGACATGCAGCCTGCCACCATGA
 GACCTCAACCCCTGCAGGACCCAGCCAAGCTGACCTTACACTCCCCACACAGAGGATGGAGGTCTCTGCCACCG
 AGAGGGCTGCTGAGGATGGAGCCTCAGTCAGCTCCAGCAGCAGAGGGCTCTGGGAGCAGGACTTCACCTTGAA
 ACCTCGGGGAGAATACGGCTGTAGTGGCGTGGAGGCCTGACCGCCGAACCAGTCCCAGTGGATCAGGGGCCAC
 GGGGGCCTCACAGGGCTCTGGACAGGAAAGAGGTGCTGGAGGGTCATTGCCGGAGGCCCTGTGGGCTCATCT
 TTGCTGTGCTGGTTCATGCTGTACCGCATGAAGAAGAAGGACGAAGGCAGCTACTCCTGGAGGAGCCG
 AAACAAGCCAACGGGGGCCCTACAGAACGCCACAAACAGGAGGAATTCTATGCCTGACGCGGGAGCCATGCC
 CCCTCCGCCCTGCCACTCACTAGGCCCCACTTGCTCTTGAAGAACTGCAGGCCCTGCCCTCCCTGCCACC
 AGGCCACCTCCCCAGCATTCCAGGCCCTCTGGCGCTCTGCCACGGAGTCGTGGTGTGGAGCTCCACTCT
 GCTTCTCTGACTTCTGCCTGGAGACTTAGGGCACAGGGTTCTCGCATAGGACCTTCCACCACAGCCAGCACCT
 GGCATCGCACCATTCTGACTCGGTTCTCCAAACTGAAGCAGCCTCTCCCAGGTCCAGCTCTGGAGGGAGGGGA
 TCCGACTGCTTGACCTAAATGGCCTCATGTGGCTGGAAGATCTGGGGTGGGGCTGGGCTCACACACCTGTAG
 CACTTACTGGTAGGACCAAGCATCTGGGGGTGGCGCTGAGTGGCAGGGACAGGAGTCACTTGTTCTGGGG
 AGGTCTAATCTAGATATGACTTGTGACATGTTCTAGTCTTGTGATAGCCCAGTAGACCTTGTGTTA
 CTTCTGAGGTAAGTAAGTAAAGTAAAGTGTGATTCGGTATCCCCCATCTGCTTCCCTAATCTATGGCGGGAGACAGCATC
 AGGGTTAAGAAGACTTTTTTTTTAAACTAGGAGAACCAAATCTGGAAAGCCAAATGTAGGCTTAGTTG
 TGTGTTGCTCTTGAGTTGTCGCTCATGTGTGCAACAGGGTATGGACTATCTGCTGGTGGCCCTTTCAAGCAGTCGTG
 TCTGGTGGCAGGCTGGCCAGTCCAGGCTGCCGTGGAGAACGCCGTGTGAGAAGTGAATGCTGGACTCAGCCT
 TCAGGGCCATGCTGAGGCCCTGGCCGTGCCACGTTGGAGAACGCCGTGTGAGAAGTGAATGCTGGACTCAGCCT
 CAGACAGAGAGGACTGTAGGGAGGGCGCAGGGCCTGGAGATCCTCTGCAGACCACXCCGTCTGCCGTGCGC
 CGTCTCCAGGGCTGCTCCCTGGAAATTGACGAGGGTGTCTGGCAGAGCTGGCTTGAGCGCCTCCATCCA
 AGGCCAGGTTCTCCGTTAGCTCCTGTGGCCCCACCCCTGGGCCCTGGGCTGGAATCAGGAATATTTCCAAAGAGTGA
 TAGTCTTTGCTTTGGCAAAACTCTACTTAATCCAATGGTTTCCCTGTACAGTAGATTTCCAAATGTAATAA
 ACTTTAATATAAAGT

FIGURE 15

GCTCCGGCCAGCCGGTCCAGAGCGCGCAGGGTCCGGGAGCTCCGCCAGGCTGGTACCTGCGTCCGCCGGC
GAGCAGGACAGGCTGCTTGTTGTGACCTCCAGGCAGGCAGGCCATCCTCTCCAGAATGAAAGATCTTCTGCCAG
TGCTGCTGGCTGCCCTCTGGGTGTGGAGCGAGCCAGCTCGCTGATGTGCTTCTCTGCTTAACCAGAACAGCAAT
CTGTACTGCCTGAAGCCGACCATCTGCTCCGACCAGGACAACACTGCGTGACTGTGCTGCTAGTGCCGGCATTGG
GAATCTCGTGACATTGGCACAGCCTGAGCAAGACCTGTCAGGCCCCTGCCCCATCCCAGAACAGCGTCAATGTTG
GTGTGGCTTCATGGCATCAGCTGCTGCCAGAGCTTCTGTGCAATTTCAGTGCGGCCATGGCGGGCTGCCGGCA
AGCGTCACCCCTGCTGGGTGCCGGCTGCTGAGCCTGCTGCCGGCCCTGCTGCCGGTTGGCCCCTGACCGCCAG
ACCCCTGCCCCGATCCCCCAGCTCAGGAAGGAAAGCCCAGCCCTTCTGGATCCCACAGTGTATGGGAGCCCCCTGA
CTCCTCACGTGCCCTGATCTGTGCCCTTGGTCCCAGGTCAAGGCCACCCCTGCACCTCACCTGCCCAAGCCCCCTGC
CTCTGCCCAAGTGGGCCAGCTGCCCTACTTCTGGGTGGATGATGTGACCTCCTGGGGACTGCCGAAGGGACG
AGGGTTCCCTGGAGTCTACGGTCAAACATCAGACCAAGTCCATGGACATGCTGACAGGGTCCCCAGGGAGACCGT
GTCAGTAGGGATGTGCGCTGGCTGTGACGTGGGTGTGCAGTGACGTGAGAGCACGTGGCGCTCTGGGGCCA
TGTTTGGGGAGGGAGGTGTGCCAGCAGCCTGGAGAGCCTCAGTCCCTGTAGCCCCCTGCCCTGGCACAGCTGCATGC
ACTTCAAGGGCAGCCTTGGGGTTCTGCCACTCCGGTCTAGGCCAAATCCAGCCAGTCCTG
CCCCAGCCCACCCCCACATTGGAGCCCTCTGCTGCTTGGCTCAAATAATACAGATGTCCCC

FIGURE 16

CAGCAGGTACAGCCCCCTCGAGGCAGACAGCGGCCCGCCGCACCAGAGCAGTGGTACAGGCATGGGATGGGAAGAAAT
 GCAGCGTATGGATGTTCTACCTCTTGATTTACTTGTAACTTCAGCTGGATTGTGGATAGTATACTTCATAGCT
 GTGGAAGATGACAAAATTTCAGCTGAAAGGAAACCTGGTGTGAAGCATGCACCATAAAGCAT
 TGCAGGTGATGATCCTCCTGCAAGCTGTGTTAGTCAAGTTATGAACATGGCAGCCTCCTAGCCCTGTGGTAG
 CTGTTCTCGCCTCATACAACGTAAACCGAAGGTTAAACCCGTGGCTGAATATTAGTGGATTGGTGGCTGTGTT
 CTGGCTCCTCGGAATGACCTTACTTGGTAATTTCAGCTCACAAATGATGAAGAAATCCATAACGTGGAACTTC
 CTTGACCTTGGATTGGCACATTGACCTGCTGGATCCAGGCTGCGCTGACACTCAAGGTCAACATCAAGAATGAAG
 GACGGAGAGTTGGAATTCCACGGGTTATTCTGTGGCATCTATCACTCTGTGTTCTACTTCATCCTCATG
 GCCCAAAGCATCCACATGTATGCAGCCAGGGTCCAGTGGGCCTGGTCATGTGCTTGTCTTATTGGCACCTT
 TGCCGTGGAGTTCCGGCATTACCGCTATGAGATTGTTGCTCTGAGTACCAAGGAGAATTCTCTAACGTTCTCAGAAA
 GCCTGTCAGAAGCTCTGAATATCAGACTGACCAGGTGTAAACCATCAGTTTCTTGCTGGTGAGGTGGGTGTA
 CAGTGGGGAGGGCCAGTAGGACACACTCACAGGACTTGACATAGAACCTCATTCACACACACACACACAC
 ATTCACTGCCACATTGCCAAATGAGCTTTCAGGGCGAGTTATTCTTTAATGAAAAAGCACAAGCCCTATGTGTT
 CGAAATACACGCTGTTACACTGAAAATATATGCACGACAGAGCAAGAAGCTGTGCATGATCACTTCTTATCCGCC
 CCTTCCCAGCACTCCCTCTCCATTCTCCACATGTCATCAAGCACCCCTACCGAGTAGGGCAGGCCAAATGTT
 CCTTGGAGTAATGCCAACTCCGACGTTGCCTCAGGTCAAAGGGCTGGAACCAGCTCGTAGGAAGTTCTGAA
 TCTGGCACTAATATTCTTGAGTGGATAATAGTGTATCATAGAATAGGACGGAAATTGTATTGAGATGTGACCTGTG
 TCGCCTGTGGAAAGGCATAGTGAGAAGAACCTTCCCACGAAAGCCCCCTCATCGTTGTTAGTGGTGGCTGTG
 GATCCCAGGAGAGACATATGCCACAGACTGTGAGAGCAAAGCCGCCGTGTGATCTGGACTTGATGCACTGTGACT
 GAGAATGATTCCAAATGTGAATATGTGTAGGGACGTGGCTATCAGGCCTGGAACAAGATGGGGCAGTGAAGGTA
 TGGTTAGTGTGTTGCTTCATAGTATGCCATGTACAATGTTATATTCTAGTTCTTTAAGTAACCTACCATGA
 GTCTCTAAGCCTCATGGACAAAGATGTAGACCAAATGCAAGAGCTGAGCTTGCCTTGGGTTCAACCATGATCAA
 GAAAAACTGAGGTACCTGCAGGCTTACGTGGAGCTAACAGACAATATC

FIGURE 17

CTGCCTCCACTGCTCTGTGCTGGGATCATGGAACCTGCACTGCTGTGGGCTGGTGGTGTGGCTGGTGTGATTCC
 AATCCAGGGCGGGATCCTGAACCTGAACAAGATGGCAAGCAAGTGACTGGGAAATGCCCATCCTCTCCTACTGGC
 CCTACGGCTGTCAGTGCAGACTAGGTGGCAGAGGCCAACCAAAGATGCCACGGACTGGTGTGCCAGACCCATGAC
 TGCTGCTATGACCACCTGAAGACCCAGGGTGCAGCATCTACAAGGACTATTACAGATAACAACCTTCCCAGGGAA
 CATCCACTGCTCTGACAAGGAAGCTGGTGTGAGCAGCAGCTGTGCCTGTGACAAGGAGGTGGCCTCTGCCTGA
 AGCGAACCTGGACACCTACCAAGCGACTGCCTTCTACTGGCGGGCCACTGCCGGGGCAGACCCCTGGTGC
TAGAAAGCCCACACCCCTCACCCCTGTTCCCTCAGCATGGAGCTCTGCATCCCCACCTCAGTATCTAACCTGAACCAGC
 CTGGCTTTCAAACACTCCGGGGGAGGTAGTCCAGCCTCCCCCGGAACCCCTTACCAATGCCTCTGACCTTCTG
 AAGCTTCCGAATCCTCCCAGTGAGGCAGTAGCTGTGTCCCTGAGGGTGGATGGAAATCTTGGGAGAAGCCAAG
 CAAGGGAGCCCTCAGAGGTGGTTGGACCAAGCATCGGGTGGGGAGGGTGTGCCGTGTCCCCACCTGCT
 GGCCCCCTGTCCCTCACCCCTCCAATATAGTCTCGAGCTACAACACTGCAGCAGCCACTATAAAGGGCAATAT
 TGATCTTCTGTCATGTGGCTCTATCTTAAACCTCAAGGCCCTCACTGTCTAACGATAAAAGCCTCTCATAGG
 CACTGGGGACCCCTGCACAGTCTGCCATGTGACCCCTCTCCCAGGCAAGCTCTGAAGTCCCTGCAGGTGGAGGCCAT
 GCCTGTCTTAAACTCAGTTGCATCCCTGGTGCCAAAGCAACACCAGAACAGAAGGAGCTCCATAATCCTTCTT
 GGGTGAAGCCTAGACAAAGCCGCCAGGTCTGTGGCTCAGGCACCAGAGCCTTGAGTACTTCTCCTGCCTCCAGG
 CATTGGCTCAGGGTGAATTACAAGGGCTACTGAATGGCTATTACTTCATCAGACTGATCCCCACCTCCTCAGGG
 TCAAAGGGCTACTTCTGGAAGTCTCCCAGGCTGACTCCTCTCCCTGACTGCAAGGGCTCACTCCCTCTCCAAG
 CTCCCACAATGCTCATGGCTCTGCCGTTACCTAGCTTGGCCTAGAGTGGCAAATGGAACCTCTGTATCTCCCC
 AACTAGACTGGAGCCCCCGAAGGATGGAGACCATGTCTGTGCCATCTGTGTTCCCTGTTTCCCACATAACTAGGT
 GCTCAATTCTGCTGTGAATGGCGTGAGCCCATAATGGATACACAGAGGTGCAGCAGATGGTGTGGGTACCTCAC
 CCAGATATCTCCAGGCCAACGCCCTCTCCCTGAGTGAGGCCAGGTGTTGGCAGCCAATGCTCCAATCTGCC
 CTTCCCTAAATACTGCCCTGGCTAGTGGAGCTGCCCTCCCCCTGCCACCTCTCCCACCAAGAGGCCACCTGT
 CACTCATGCCAGGAGAGTGACACCAGGAGGTACAATTGCCAGCTCCCCGTGTGTGCAGCAGATTGTCTGGGTT
 GAATGACACTCTCAAATTGTTCTGGATGGCTGAGGCCAGGCCTCTCCTGGAACCACCTCTGCTTGGTCTGA
 CCCCTGGCCTATCCAGTTCTGGTCCCTCACAGGTTCTCCAGAAAGTACTCCCTCAGTAAAGCATTGACA
 AGAAAAAAAAAAA

FIGURE 18A

GGCTGAAAGAGCCTGAGCTGTGCCCTCCATTCCACTGCTGTGGCAGGGTCAGAAATCTGGATAGAGAAAACCTT
 TGCAACCGGAATGTATCTTGTAAATCCCTAGCACGAAAGACTCTAACAGGTGTGCTGTGGCCAGTTCACCAACCA
 GCATATCCCCCTGCCAAGTGCACACCCCAGCAAAATGAAGAGGAAGCAAAACAGGTGGAGACTCAGCCTGAGA
 AATGGTCTGTCAGCACACCCAGAGCTACCCAACAGATTCTATGGAGTCTGAATTCCAGGGTGGCGGATAT
 TCCAATAAAGCC**ATG**TATATCCGTGTACCTATGACACCAAGCCAGACTCACTGCTCCATCTCATGGTAAAGATTG
 GCAGCTGGAACCTCCCAAGCTTAATATCTGTGCATGGAGGCCTCCAGAACTTGTAGATGCAGCCAAAGCTGAAAC
 AAGTCTTGGAAAGGCCTGATCAAGGCTGCTATGACCACCGGGCTGGATCTCACCGGGGTGTCAGCACAGGT
 GTTATCAGCCACGTAGGGATGCCCTGAAAGACCACTCCTCAAGTCAGAGGCCGGTTGTGCTATAGGAATTGC
 TCCATGGGCATCGTGGAGAATAAGGAAGACCTGGTGGAAAGGATGTAACAAGAGTGACAGACCATGTCCAACC
 CTCTAAGTAAGCTCTGTGCTCAACAACCTCCACACCCACTTCATCTGGCTGACAATGGCACCCCTGGCAAGTAT
 GGCGCCGAGGTGAAGCTGCGAAGGCTGCTGGAAAAGCACATCTCCCTCAGAAAGATCAACACAAGACTGGGGCAGGG
 CGTGCCCTCGTGGGTCTCGTGGAGGGGGCCCTAACGTGGTGTCCATCGTCTTGGAAATACCTGCAAGAAGAGC
 CTCCCATCCCTGTGGTGAATTGTGATGGCAGCGACGTGCCGACATCTGTCTTGGCAGAACAGTACTGTGAA
 GAAGGCGGAATAATAATGAGTCCCTCAGGGAGCAGCTCTAGTTACGATCAGAAAACATTAAATTATAAAGGC
 ACAATCACATCAGCTGTTGCAATTATAATGGAGTGCATGAAGAAGAAGAAACTCGTCACTGTGTTGAGATGGTT
 CTGAGGGCCAGCAGGACATCGAGATGGAATTAACTGCCCTGCTGAAAGGAACAAACGTATCTGCTCCAGATCAG
 CTGAGCTTGGCAGCGCTGGAAACCGCGTGGACATAGCACCAAGCCAGATCTTGTCTTGGGCCCCACTGGACGCC
 CCTGGGAAGCCTGGCACCCCCGACGGACAGCAAGCCACGGAGAAGGAGAAGAACCCATGCCACCAAGG
 GAGGAAGAGGAAAGGGAAAGGCAAGAAGAAGGGAAAGTGAAGAGAGAAGTGGAGGAAGAAACTGACCCCCGGAAG
 ATAGAGCTGCTGAACCTGGGTAATGCTTGGAGCAAGCGATGCTAGATGCTTAGTCTAGATCGTGTGACTTGT
 GAAGCTCTGATTGAAAACGGAGTGAACATGCAACACTTTGACCATTCGAGGCTGGAGGAGCTATAACACAA
 GACTGGGCCACCAAACACACTCATCTGCTGGTGGAGGATGTGAAAAAGAGCAACCTCCGCTGATTACCACATC
 AGCCTCATAGACATCGGGCTCGTGGAGTACCTCATGGGAGGAGCCTACCGCTGCAACTACACTCGGAAAACCT
 TCGGACCCCTTACAACAACCTGTTGGACCAAAGAGGCTAAAGCTCTAAACTCTGGGAATGGAAGATGATGAGC
 CTCCAGCTAAAGGAAAGAAAAAAAAAGAAAAAGGAGGAAGAGATGACATTGATGTGGACGACCTGCCGTG
 AGTCGGTTCCAGTATCCCTCCACGAGCTGATGGTGTGGCAGTGTGATGAAACGCCAGAAAATGGCAGTGTCC
 CTGGCAGCGAGGGGAAGAGAGCATGGCCAAGGCCCTGGTGGCCTGCAAGCTCTACAAGGCCATGCCACGAGTCC
 CCGAGAGTGTCTGGGATGACATCTCCAGGACTGGATAACAATTCCAAGACTTCGGCCAGCTTGTCTTGGAG
 TTATTAGACCAGTCTATAAGCATGACGAGCAGATCGCTATGAAACTCTGACCTACGAGCTGAAAACACTGGAGCAA
 CTCGACCTGCCCTCAAACCTGGCGTGGCAGCAGGAAACACCGGACTCTATTGTCACACCTGCAAGCCAGATGCTG
 CCGATATGTGGATGGGAAGACTGCGGATGCGGAAGAACCCCCGCTGAAGGTTATCATGGGATTCTCTACCCCC
 ACCATCTGTTTGGAAATTGCGCACATATGATGATTCTGCTATCAAACATCCAAGGAAACGAGGATGGCAAAGA
 AAAAGAAGAGGAAATAACGGATGCAATGAGATGCTGGCTCAAGAAAGGGGATGAGGAGAACGAGCATAAAAAC
 AGAGAAGTATTCCATCGGAACAAAGATCTGTGAATTCTATAACGCCCTATTGCAAGTTCTGGTTACACAATA
 TCATACTGGGCTACCTGCTGTTAACTACGTCATCCTGGTGGGATGGATGGCTGGCCCTCCAGGAGTG
 GATCGTCATCTCTACATCGTGAGCCTGGCTTAGAGAAGAGATACGAGAGATCCTCATGTCAGAACCCAGGCAAAC
 GCCAGAAAATCAAAGTTGGCTCAGGAGTACTGGAACATCACAGATCTGTCGGCCATTCCACATTGATTGGA
 GCAATTCTCGCCTACAGAACCCAGCCCTACATGGGCTATGGCGGGTGTACTACTGTGTGGATATCATCTGGTA
 CATCGTCTGACATCTTGGTGTCAACAAGTATCTGGGCCATACGTGATGATGATTGGAAAGATGATGATCG
 ACATGCTGTACTTGTGGTACATGCTGGTGTGCTCATGAGTTGGAGTAGCCCGTCAAGCCATTCTGCATCCA
 GAGGAGAAGCCCTCTGGAAACTGGCCGAAACATCTTCTACATGCCCTACTGGATGATCTATGGAGAGGTGTTGC
 AGACCAAGATAGACCTCTACGCCATGAAATTAACTCTCTTGTGGTGGAGAACCTATATGATGAGGAGGGCAAGCC
 TTCCCTCCCTGTATCCCCGGCCCTGGCTCACTCCAGCACTCATGGCGTGTATCTACTGGTCGCCAACATCCTG
 GTGAAACCTGCTGATGCTGTTCAACAAACTTCTTGAAGTAAAATCAAATCTAACCCAGGTGTGGAAGTCCA
 GCGATATCAGCTGATTATGACATTCTCATGACAGGCCAGTCCCTGCCCCCACCAGATGATCATTTAAC
 TCATCATTATGCGTCTCAGCGGCCGCTGCAGGAAAAGAGAGAAGGGACCAAGAGGAACGGGATCGTGGATTGAAG
 CTCTTCTTAGCGACGAGGAGCTAAAGAGGCTGATGAGTCAGGAGCAGTGCCTGAGGAGCACTTCCGGGAGAA
 GGAGGATGAGCAGTCAGCTCAGCGACGAGCGATCCGGTCACTCTGAAAGAGTGTGAAAATATGCAATGAGGT
 TGGAGAAATCAATGAAAGAGAAACTTTATGAAAATCTCGGGGAATGACAGGTGACCTGACTGCTCAGCTAGAA
 GAATTATCTAACAGAACGGTGAATGCTCTGAAAATCTCGGGGAATGACAGGTGACCTGACTGCTCAGCTAGAA
 CCGGGCTTCTCTGAATGAGGCAACGTATCTCTCCGGCAAAGCAGCATCAATAGCGTGTGGCTACAGCTTGT
 ATCGATATCATTAAACGGAGAAGAGTTATTGAGGATACATCTCTCCACGTCACCAGGGACAG

FIGURE 18B

GAGTCAGGAAAAACCTGTCCTCCGTATAAGGAAGAGAAGGCACGTGAAACGCACCTAGTCCCAGAATGTCAG
AACAGTCTCACCTTCACTGGCACAAGCACATCAGCAACCCAGATGGCAGTCACCTTGCACTAGATGACTTAAA
GAACGCTGAAGAGTCAAATAGGTCCAGATATTGGGATTCAAAGGAAGATGATGAAAGACAGACAGACTCTAAA
AAGAAGAAACTATTCCTCCAAGTTAAATAAAACAGATGTGATACATGGACAGGACAATCAGATGTTCAAAACACT
CAGCTAACAGTGGAAACGACAATATAGAAGGCACATTTCTATCCCTGGAAGAAACAAAATTACACGCTATT
CCCCGATGAAACGATCAATGTTGTAACAAATGAAGTCCAGAAGCTCGTCTATTCCCGGGAAAGAAAGCTGGTCG
GTGGGGTTAACCAGGATGTAGAGTACAGTTCAATCACGGACCAGCAATTGACGACGGAATGGCAATGCCAAGTTCAA
AAGATCACGCGCTCATAGCACAGATATTCTTACATTGTCGGAAGCTGCAGTGCAAGCTGAGCAAAAGAGCA
GTTGCAAGATATGCAAGATGAACACCATGTCGCTGAAGCAATTCTCGAATCCCTCGTTGCCCTAACCTACTG
ACAGAAATGGGATGGAAAACCTACTGTCGCTGAAGCCAGATCAAACCTTGGGATTCCCATCTCAGGTCAAAAGT
TTACATGGACATCCTAGGAATGTGAAATCCATTCAAGGAAAGTTAGACAGATCTGGACATGCCAGTAGTGTAAGCAG
CTTAGTAATTGTCGCTGGAATGACAGCAGAAGAAAAAAAGGTTAAGAAAGAGAAAGCTCCACAGAAACTGAATGCT
AGTCTGTTTCTTTAATTTTTAAACAGTCAGAAACCCACTAATGGGTGTCATCTGGCCATCCTAAA
CACATMTCCAATTCTAAACATTCCCTAAAAAATTTGGAAATTCAAGACTTGATTTACAATTAAATGCACT
AAAAGTAGTATTTGTTAGXATATGTTAGTAGGCTTAGTTTCAGTTGCAGTAGTATCAAATGAAAGTGATGATA
CTGTAACGAAGATAATTGGCTAATCAGTATAAAGATTATAACATCTTTATTACTGAGGGCCACCAAATAGCCT
AGGAAGTGCCCTCGAGCACTGAAGTCACCATTAGGTCACTCAAGAAGTAAGCAACTAGCTGGGCACAGTGGCTCATG
CCTGTAATCCTAGCACTTGGGAGGCCAAGGCAGAAAGATAGCTGAGTCCAGGAGTTGAGACCAGCCTGGCAAC
ATAGTGATACCCATCTTTAAAAAAAAAAAAAA

FIGURE 19

CTGAATCTCGTTCTCTCCAGGGACCCCTCATTTCATATCCAGAAAATGTGATGCCAACAGGTATCAGCGT
 CTGGATGCCACTTCACGTTAGCCACAAGTGACTCAGTGGAAAGATCCAGAGTCACAGAGGCTCGTCAGGAAG**AT**
GTCTACAGAAAAGGTAGACCAAAAGGAGGAAGCTGGGGAAAAAGAGGTGTGCGGAGACCAGATCAAAGGACCGGACA
 AAGAGGAGGAACCACCAGCTGCTGCATCCCAGGCCAGGGTGGCGTCCAGGTGGCAGAGCAGCTAGGAACGCAAGG
 CCTGAACCTGGGGCCAGACACCCCTGCTCCCGGCCATGGTCACGACCCCTCCAGTACCTGCCTACTGTGGGCCA
 GGAGGTGGCCAAGTCTTGGCAGGCCGTGCCCCAGGCTGCTGCAGTTGGGGTGTCTTCTGCACCATCTCC
 TTTTGCTCTGGGTGTCTGCTTCCTCTATGGCTCCTTACTATTCCATATGCCGACAGTCAGCCACCTCAGCCCT
 GTGCATTCTACTACAGGACCGACTGTGATTCCCTACCACCTCACTCTGCTCCTCCCTGTTGCCAATGTCTCGCT
 GACTAAGGGTGGACGTGATGGGTGCTGATGTATGGACAGCCGTACCTGTTAGAGCTTGAGCTGCCAGAGT
 CCCCTGTGAATCAAGATTGGCATGTTCTGGTCACCATTCCTGCTACACCAGAGGTGGCGAATCATCTCCACT
 TCTTCGCGTTCGGTGATGCTGCATTACCGCTCAGACCTGCTCCAGATGCTGGACACACTGGCTTCTCTAGCCTCCT
 GCTATTGGCTTGCAGAGCAGAAGCAGCTGCTGGAGGTGGAACCTACGCAGACTATAGAGAGAACTCGTACGTGC
 CGACCACTGGAGCGATCATTGAGATCCACAGCAAGCGCATCCAGCTGTATGGAGCCTACCTCCGCATCCACGCGCAC
 TTCACTGGGCTCAGATACTGCTATAACACTCCGATGACCTGCGCCTCATAGGTGTTGCCAGCAACTCACCT
 CCTCAGCGTCATCGTCTTCAGCTACATGCAGTGGGTGCCCCCGACACCGCTTCTTTGC
 AGGTTAACATCCAAAAAGAGACAATTCCCGAAGGAAGTCCAACGAAGGATCTGCTCATGCCAGGGCTGAA
 GGCCAGGAGGAGTCAACTCCGCAATCAGATGTTACAGAGGATGGTGAGAGCCCTGAAGATCCCTCAGGGACAGAGGG
 TCAGCTGTCCGAGGAGGAGAAACCAGATCAGCAGCCCCCTGAGCGGAGAAAGAGGAGCTAGAGCCTGAGGCCAGTGATG
 GTTCAGGCTCCTGGGAAGATGCAGCTTGCTGACGGAGGCCAACCTGCCTGCTCCTGCTCCTGCTTCTGCC
 CCTGCTCTAGAGACTCTGGGAGCTGAAACCTGCTGGGGTGTCTCCGACAGCGCCCCACCTGCTCTAGTTCC**TG**
AAGAAAAGGGCAGACTCCTCACATTCCAGCACTTCCCACCTGACTCCTCTCCCTGTTTCTTAATAACT
 ATTTGTGTCAGCTCAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 20A

AACCGGCTGGGGGATGGGGCCACCGCTCCGCTGCTGCTGCTACTGCTGCTGCCGCCACGCCTCGCCT
 GCCCCCCCTCGCTCGTCCCCCGCGCCGGCAGCTCCCAGGGCTCTGGCTGCCCTGCTGAGGAGGGCCTCGCG
 AGCGTCCGAGGCCGTGTGAACGATGGAGTGTGGAAAGGTCAGAACGGTTCCGGCAATGGACTTTACCGCTACG
 AGGTGTCGCCGTGGCCCTGCAGCGCTGCGCTGGCGTTGAGAACGCTTCCGGCACAGGTTACGTGGCAGGAT
 GACTATACTCAGTATGTGATGGACCAGGAACCTGCAGACCTCCGAAAACACTACCTGAGGCCCTGAAGCATCCAG
 CCCAGCCAGGCCCTCAAAACACAGCGTTGGCAGCGAGAGGAGGTACAGTCGGAGGGCGGTGCTGCCCTGGCAACG
 CCCTCCGACGCCACCTGCCCTCTGGAGGCCCTGCTCCAGGCCCTCAGACGTGCTGCCAGGACCCATACG
 GCGCAGGACAGACCCCCCGCTGAGGGTATGACCGCTCTCCGAGAGACATCCTGACCTATGTGGCCACACGTCTGC
 GCTGACGTACCCCTCCGGGCCCGAACCGCTCCGGAGGACCTCCCTGCCGCCAGCTCCAGGCCAG
 ATGAGCTCAGCCCTAAGGTGGACAGTGGTGTGGACAGACACCATCTGATGGCGCCCTCAGTGCCTATGCTGCCAG
 AGGCCCGAGCTCCCCCGGGGAGGGCAGCCTGGAGGCCACAGTACCTCTGCGTGCACCCCTCAAGAATGCCAGGCC
 TTTGCTGGCACCGCCGGGGGGAGGGAGGGAGCTGGCCTCACCTCTGGAGATTCCGAAGACCCCTCCAGCACAGGCGATG
 GAGCACGGATTCATACCCCTCTGAAGGACCTGAGGCCCTGAGGGTGGCTGAGGTGAGGGGCTGAGTGGCTGGAG
 GACGGCATGGCTGAGCTGATGGCTGGCGTGTGACAGGCCAGGGCTGGAGATGGAGTAGCTGAGGCCAGCCCTGGAG
 GGCCCTGGGAGAGTCTGGAGAACAGGGGATGGCCCAAGGCCACCCCTCGTGGAGACAGCTTCCAGATGACGGAG
 TGCAGGACGACGATGATAGACTTACCAAGAGGTCACCGCTGAGTGCCACACTCGGGGGCTCCTGAGGCCAC
 GGGTCTGACTCTTACCTGGAGCCCTCCCTTGCAGGCCCCCTGACATGGAGAGGAAGAAGTCCGAGCACCCCTGA
 GTCTTCCCTGCTCTCAGAAGAGGAGACTGCCAGTGGAGAACGTCAAGAGCCAGACGTATTCAAAGATCTGCTGG
 GGCAGCAGCCGATTGGAGCCGGGGCGCTGGCTGGAGCTCCAAAGGCCACCCCTCGTGGAGACAGCTTCCAGATGACGGAG
 GAGCAGAGCCTCCAGGGGTGCTCAGGAGGCCCTCAGCGACGGCTGCAATTGGAGGTCCAGCCTCCAGGAAGAAG
 GGCGGGGGTACATCGTACAGACAGAGACCCCTGCGCCCCGAGGAAGGAAGGGCGCTGGTGGAGGACGTGCC
 GCCTCCTGCAGGTGCCAGCAGTGCCTCGTGCAGTGGAGGTCTGGACCAGCAGTGCACCTCAAAGTGAGGCC
 AATGTCCAAAACGTGACCACTGAGGATGTGGAGAACGCCACAGTTGACAACAAAGACAAACTGGAGGAACCTCTGG
 ACTGAAAATTCTCAAACCGAGTCGGGTGAAAGAACAAACTCAAGTTCTGCCTCCTCAGGCCAGCAGCGCTCATCTGG
 CCACCAAGTTCATCGCCTCACCTGGTCTCCCTGCCTGCATCCTGGCGTCTGGCCTCTGGCCTCATCTAC
 TGCCTCCGCCATAGCTCTCAGCACAGGCTGAAGGAGAACGCTCTGGGACTAGGGGCGACCCAGGTGAGATGCCAC
 TGCCGCCTACCAGGAGCTGCGCCAGCGTATGGCACGCCAGCAGGACCTGAGGGCCCGCACACGTAC
 GCATCAGCAGCGTCTCATCCAGTTGAGCAGGCCGATCCCCAGGCCCTCCGACAGCAGCGCTCATCTGG
 TCCGAGGAGCTGTGAGTCCAACATGGACATCTCCACCGGCCACATGATCTGTCTACATGGAGGACCTGAA
 GAACAAGAACCGGCTGGAGAACGGAGTGGAGGCCCTGCGCCCTACCGGGGAGGCCAACAGCTCGTCTGGGCC
 AGAGGGAGGAGAACGTGCCAAGAACCGCTCCCTGGCTGTGACCTATGACCATCTCCGGTCTGTGAAGGCG
 GAGAACGCCACGCCACTCAGACTACATCAACGCTAGCCCCATCATGGATCACGACCCGAGGAACCCCGTACAT
 CGCCACCCAGGGACCGCTGCCACCGTGGCTGACTTTGGCAGATGGTGTGGAGAGCGGCTGCGTGGTATCG
 TCATGCTGACACCCCTCGCGAGAACGGCGTCCGGCAGTGCCTACACTGGCCGGATGAAGGCTCAAATCTTAC
 CACATCTATGAGGTGAACCTGGTCTCCGAGCACATCTGGTGTGAGGACTTCTGGTGTGAGGAGCTCTATCTGAAGAA
 CCTGCAGACCAACCGAGACGCCACCGTGCAGCAGTCCACTCTGAGTTGGTATGACCGAGGAGTCCCTCCTCCT
 CAAGGTCCCTCTGGACTTCCGAGAAAAGTAAACAAATGCTACAGGGGGCTTGTGCGATAATTGTTATTGC
 AGTGACGGTGCAGGCCGGAGCGGCACCTACGTCTGATCGACATGGTCTCAACAAAGATGCCAAAGGTGCTAAAGA
 GATTGATATCGCAGCGACCCCTGGAGCACCTGAGGGACCAGAGACCCGGCATGGTCCAGACGAAGGAGCAGTTGAGT
 TCGCGCTGACAGCGTGGCTGAGGAGGTGAACGCCATCTCAAGGCCCTCCAGTTGAGCGGCAGCGTCAGGGGCC
 TCAGGGGAGCCCCACCCACGGATGTGTGAGGAATCATGATCTGACTTTAATTGTGTCTTCTATTATAACTGC
 ATAGTAATAGGGCCCTTAGCTCTCCGTTAGTCAGCGCAGTTAGCAGTTAAAAGTGTATTGGTTAATCAAACAA
 TAATAAGAGAGATTGTGGAAAAATCAGTTACGGGTGGAGGGGAATCGGTTCATCAATTTCACTTGCTTAA
 AAATACTTTCTTAAAGCACCGGTCACCTCTGGTTGAAGTGTGTTAACATGCAGTAGCCAGCACGTTCCAG
 GCGGTTCCAGGAAGAGTGTGCTTGTCACTGCCACTTCCGGAGGGTGGATCCACTGTGCAAGGAGTGGCCGGGAA
 GCTGGCAGCACTCAGTGAGGGCGCCGGCACACAAGGCACGTTGGCATTCTCTTGGAGAGAGTTATCATGGGA
 GAAGCCGGGGAGACAAGTGAACGTCTGCAGCTCCGGCAAGTGAAGACAATCACAGCTCTCGCTGGTCTCCA
 TCAACACTCGCCGGGTACCATGGACGGCCCGTCAAGCCACACGGTCAAGCCAGCACAGAGTGTATTCAAGGGCTCC
 CGGGGGCAGACACCTGTGACCCCATGAGTAGTGTGCCCACCTGAGGCTGGCACTCCCTGACCTCACCTTGCAAAGT
 TACAGATGACCCCCAACATTGAGATGTGTTTAATGTTAAATATTGATTCTACGTTATGAAAACAGATGCC
 GTGAATGCTTACCTGTGAGATAACCACACCAGGAAGAACAAATCTGGCATTGAGCAAGCTATGAGGGTCCCCGGG
 AGCACACGAACCTGCCAGGGCCCCGCTGGCTCCTCCAGGCACGTCCGGACCTGTGGGCCAGA

FIGURE 20B

GGGGACATTCCTCCTGGAGAGAAGGAGATCAGGGCAACTCGGAGAGGGCTGCGAGCATTCCCTCCGGAGAG
GAAATCAGGGCGACCTGCACGCAGTGCCTGGAGGGAACTGAGAAACCAGCCGACCGGCCCTGCCCTCT
TCCCAGGATCACTTAATGAACCACGTGTTTGACATCATGTTAACCTAACGACGTACAGATGATTCCGGATTGACA
AAATAACATTTGAGTATCCGATTGCCATCACCCCTACCCCCGAAATAGGACAACTCACTTCATGACCAGGATGAT
CACATGGAAGGCAGCGCAGAGGCAGCTGTGGCTGCAGATTCTGTGGCTCAGCGTATAAACGCACCTC
CATCCCGCCCTTCCCACAGCATTCCATCTTAGATAGATGGTACTCTCAAAGGCCCTACCAGAGGAAACACGGC
CTACTGAGCGGACAGAATGATGCCAAATATTGCTTATGCTCTACATGGTATTGTAATGAATATCTGCTTTAATAT
AGCTATCATTCTTTCCAAAATTACTCTCTTATCTGGATTAAATTGAAATGAAATTATCTGAATATAGG
AAGCATATGCCTACTTGTAAATTCTAACTTATGTTGAAGAGAAACCTCCGGTGTGAGATATAACAAATATATT
AATTGTGTCATATTAAACTTCCCGGAATTC

FIGURE 21

GCATCTGGTTGTCAGATCCGAGAGGCTCTGAAACTGCGGAGCGGCCACCGGACGCCCTCTGGAGCAGGTAGCAGCA
TGCAGCCGCCTCCAAGTCTGTGCGGACGCCCTGGTTGCGCTGGTTCTGCCTGCGGCCTGTCGGATCTGGGA
GAGGAGAGAGGCTCCCGCCTGACAGGGCCACTCCGCTTTGCAAACCGCAGAGATAATGACGCCACCCACTAAGAC
CTTATGCCCAAGGGTCCAACGCCAGTCTGGCGCGTCGTTGGCACCTGCGGAGGTGCCTAAAGGAGACAGGACGG
CAGGATCTCGCCACGCACCCTCCCCCTCCCCGTGCCAAGGACCCATCGAGATCAAGGAGACTTCAAATACATC
AACACGGTTGTCTGCCTGTGTTCTGCTGGGATCATCGGAACCTCCACACTCTGAGAATTATCTACAAGAA
CAAGTGCATGCGAACGGTCCAATATCTTGATGCCAGCTGGCTCTGGGAGACCTGCTGCACATGTCATTGACA
TCCCTATCAATGTCTACAAGCTGCTGGCAGAGGACTGCCATTGGAGCTGAGATGTGTAAGCTGGTGCCTTCATA
CAGAAAGCCTCCGTGGAATCACTGTGCTGAGTCTATGTGCTTGAGTATTGACAGATATCGAGCTGTTGCTTCTG
GAGTAGAATTAAAGGAATTGGGTTCCAAAATGGACAGCAGTAGAAATTGTTGATTGGTGGTCTCTGTGGTTC
TGGCTGTCCTGAAGCCATAGGTTTGATATAATTACGATGGACTACAAAGGAAGTTATCTGCGAATCTGCTTGCTT
CATCCCGTTCAGAAGACAGCTTCATGCAGTTTACAAGACAGCAAAGATTGGTGGCTGTCAGTTCTATTCTG
CTTGCCATTGCCATCACTGCATTTTATACACTAATGACCTGTGAAATGTTGAGAAAGAAAAGTGGCATGCGA
TTGCTTTAAATGATCACCTAAAGCAGAGACGGGAAGTGGCCAAACCGTCTTGCGCTGGCCTTGTCTTGCCCTC
TGCTGGCTCCCCTCACCTCAGCAGGATTCTGAAGCTCACTCTTATAATCAGAATGATCCAATAGATGTGAACT
TTTGAGCTTCTGTTGGTATTGGACTATATTGGTATCAACATGGCTTCACTGAATTCTGCATTAACCAATTGCTC
TGTATTGGTGGAGAAAAGATTCAAAACTGCTTAAAGTCAAGCTAATGATCACGGATATGACAACCTCCGTTCCAGTAA
CAGTCCTGGAGGAAAAGCAGTCGTGCTTAAAGTCAAGCTAATGATCACGGATATGACAACCTCCGTTCCAGTAA
TAAATACAGCTCATCTTGAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MKENYCLQAALVCLGMLCHSHAFAPERRGHLRPSFHGHHEKGKEGQVLQRSKRGWVNQFFVIEEYTGPDPVLVGRL
 HSDIDSGDGNIKYIISGEAGTIFVIDDKSGNIHATKTLDRERAQYTLMAQAVDRDTNRPLEPPSEFIVKVQDIND
 NPPEFLHETYHANVPERSNVGTSVIQVTASDADDPTYGNSAKLVYSILEGQPYFSVEAQTAGIIIRTALPNMDREAKEE
 YHVVIQAKDMGGHMGGLSGTKVTITLTDVNDNPKFPQRILYQMSVSEAAVPGEEVGRVKAKDPDIGENGLVTYNIV
 DGDGMESFEITTDYETQEGVIKLKKPVDTERAYSLKVEAANVHIDPKFISNGPFKDTVTVKISVEDADEPPMFLA
 PSYIHEVQENAAAGTVVGRVHAKDPDAANSPIRYSIDRHTDLDRFFTINPEDGFIKTTKPLDREETAWLNTVFAAE
 IHNRHQEAQPVVAIRVLDVNDNAPKFAAPYEGFICESDQTPLSNQPIVTISADDKDDTANGPRFIFSLPPEIIHNP
 NFTVRDNRDNTAGVYARRGGFSRQKQDLYLLPIVISDGGIPPMSSNTLTIKVCGCDVNGALLSCNAEAYILNAGLS
 TGALIAILACIVILLVIVVLFTLRRQKKEPLIVFEEEDVRENIITYDDEGGGEEDTEAFDIATLQNPDGINGFIPR
 KDIKPEYQYMPRPGLRPAPNSVDVDDFINTRIQEADNDPTAPPYDSIQIYGYEGRGSVAGSLSSLESATTDSLDDYD
 YLQNWGPRFKKLADLYGSKDTFDDDS

Signal sequence.

amino acids 1-22

Transmembrane domain.

amino acids 617-637

N-glycosylation sites.

amino acids 455-458, 540-543

Glycosaminoglycan attachment sites.

amino acids 83-86, 93-96

N-myristoylation sites.

amino acids 108-113, 215-220, 242-247, 246-251, 247-252, 399-404, 594-599, 599-604, 614-619, 618-623, 749-754, 753-758, 787-792

Cadherins extracellular repeated domain signatures.

amino acids 147-157, 256-266, 476-486

Cadherin cytoplasmic region.

amino acids 641-789

Cadherin domains.

amino acids 59-150, 164-259, 273-375, 388-479, 492-593

FIGURE 23

MYGKSSTRAVLLLGIQLTALWPIAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTWTWNFRPLDGGPEQFVF
YYHIDPFQPMMSGRFKDRVSDGNPERYDASILLWKLQFDDNGTYTCQVKNPPDVDGVIGEIRLSVVHTVRFSEIHFL
ALAIGSACALMIIIVVVLFQHYRKKRWAERAHKVVEIKSKEEERLNQEKKVSVYLEDTD

Signal sequence.

amino acids 1-21

N-glycosylation sites.

amino acids 39-42, 118-121

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 205-208

N-myristoylation sites.

amino acids 15-20, 119-124, 159-164

Immunoglobulin domain.

amino acids 40-125

FIGURE 24

MAEQQQGRELEAECPVCWNPFNNTFHTPKMLDCCHSFCVECLAHLSLVTPARRLLCPLCRQPTVLASGQPVTDLPTD
TAMLTLRLEPHHVILEGHQLCLKDQPKSRYFLRQPRVYTLSDLGPQPQGGQTGPPPDTASATVSTPILIPSHHSLREC
FRNPQFRIFAYLMAVILSVTLLIIFSIFWTKQFLWGVG

Transmembrane domain.

amino acids 162-182

N-glycosylation site.

amino acids 21-24

N-myristoylation site.

amino acids 68-73

Zinc finger, C3HC4 type, signature.

amino acids 32-41

FIGURE 25

MATAAGATYFQRGSLFWFTVITLSFGYYTWVVFWPQSIPYQNLGPLGPFTQYLVDHHHTLLCNGYWLA
LVIHVGESLYAIALCKHKGITSGRAQLLWFLQTFFFGIASLTILIAKRKRQKQT

Transmembrane domain.

amino acids 57-77, 92-112

N-myristoylation site.

amino acids 87-92

Leucine zipper pattern.

amino acids 46-67

FIGURE 26

MASKIGSRRWMLQLIMQLGSVLLTRCPFWGCFSQLMLYAERAEARRKPDIPVPYLYFDMGAAVLCASFMSFGVKRRWFALGAALQLAISTYAAYIGGYVHYGDWLKVRMYSRTVAIIGGFLVLASGAGELYRRKPRSRSLQSTGQVFLGIYLICVAYSLQHSKEDRLAYLNHLPGGELMIQLFFVLYGILALAFLSGYYVTLAAQILAVLLPPVMLLIDGNVAYWHNTRVEFWNQMPLLGESEVGIFGTAVILATDG

Transmembrane domains.

amino acids 6-26, 51-71, 106-126, 138-158, 174-194, 201-221, 237-256

Glycosaminoglycan attachment site.

amino acids 125-128

N-myristoylation site.

amino acids 245-250

FIGURE 27

GRGSPLALLIRMKTLFGVWALLALILCPGVPEELFEVSIWPSQALVEFGQSLVCNCSTTCPDPGPMSGIETFLKKTQ
VDKGPKWKEFILLEDTENSILQCFFSCAGIQKDTSLGITVYQPPEQVILELQPAWVAVDEAFTVKCHVPSVAPLES
TLALLQGNQELHRKNFTSLAVASQRAEVIISVRAQKENDRCNSSCHAELDLSLQGGRLFQGSSPIRIVRIFEFSQSP
HIWVSSLLEAGMAETVSCEVARVFPKAKEVMFHMFLSQELSSLSWEGDTAWANATIRTMEAGDQELSFCASLGAME
QKTRKLVHSYSFPPPILELKESYPLAGTDINVTCGHVLTPSPSTLRLQGAPDLPAGEPAWLLLTAREEDDGXNFSC
EASLVVQGQRLMKTIVIQLHILKPQLEESSCPGKQTWLEGMEHTLACVPKGNPAPALVCTWNGVVFDLEVPQKAT

Signal sequence.

amino acids 1-30

N-glycosylation sites.

amino acids 56-59, 169-172, 196-199, 285-288, 339-342, 382-385

N-myristoylation sites.

amino acids 3-8, 242-247, 335-340, 380-385, 425-430

Intercellular adhesion molecule (ICAM) homology.

amino acids 17-123

FIGURE 28

MLPRLLLICAPLCEPAELFLIASPSPTEGSPVTLCKMPFLQSSDAQFQFCFFRDTRALGPGWSSSPKLQIAAMW
KEDTGSYWCEAQTMASKVLRSSRQINVHIPVSRPILMLRAPRAQAAVEDVLELHCEALRGSPPILYWFYHEDITLG
SRSAPSGGGASFNLSTEEHSGNYSCEANGLGAQRSEAVTLNFTVPTGARSNHLTSGVIEGLLSTLGPATVALLFC
YGLKRKIGRRSARDPLRSLPALPQEFTYLNNSPTPGQLQPIYENVNVVSGDEVYSLAYYNQPEQESVAAETLGTHMED
KVSLDIYSRLRKANITDVEDAM

Signal sequence.

amino acids 1-16

Transmembrane domain.

amino acids 211-231

N-glycosylation sites.

amino acids 167-170, 177-180, 197-200, 322-325

Glycosaminoglycan attachment site.

amino acids 160-163

N-myristoylation sites.

amino acids 31-36, 82-87, 161-166, 163-168, 176-181, 187-192, 203-208, 212-217,
216-221

Amidation site.

amino acids 238-241

Immunoglobulin domains.

amino acids 31-88, 126-182

FIGURE 29

MLLWASILLAFAFAPVCGQSAAAHKPVISVHPPWTTFFKGERVLTLCNGFQFYATEKTTWYHRHYWGEKLTLPGNLEV
 RESGLYRCQARGSPRSNPVRLLFSSDSLILQAPYSVFEGDTLVLRCHRRKEKLTAVKYTWNNGNILSISNKSWDLLI
 PQASSNNNGNYRCIGYGDENDVFRSNFKIIKIQELFPHPPELKATDSQPTEGNSVNLSCTQLPPERSDTPLHFNF
 DGEVILSDWSTYPELQLPTVWRENSGSYWCGAETVRGNIHKHSPSLQIHVQRIPVSGVLLETQPSGGQAVEGEMLVL
 VCSVAEGTGDTTFWSHREDMQESLGRKTQRSLRAELELPAIRQSHAGGYYCTADNSYGPVQSMVLNVTVRETPGNRD
 GLVAAGATGGLLSALLLAVALLFHCWRRRKSGVGFLGDETRLPPAPGPGESSHICPAQVELQSLYVDVHPKKGDLV
 YSEIQTTQLGEEEANTSRTLLEDKDVSVVYSEVKTQHPDNSAGKISSKDEES

Signal sequence.

amino acids 1-15

Transmembrane domain.

amino acids 387-407

N-glycosylation sites.

amino acids 147-150, 209-212, 374-377, 478-481

Glycosaminoglycan attachment site.

amino acids 416-419

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 413-416

Tyrosine kinase phosphorylation site.

amino acids 457-463

N-myristoylation sites.

amino acids 15-20, 81-86, 89-94, 140-145, 163-168, 205-210, 257-262, 315-320,
 355-360, 382-387, 386-391, 391-396, 394-399, 395-400

Amidation site.

amino acids 332-335

Immunoglobulin domains.

amino acids 37-87, 116-169, 205-263, 303-361

FIGURE 30

MTVIRFFPAASATKRVLPPLRVSSPRTWNPNVPESPRI PAPRLPKRMSGAPTAGAALMLCAATAVLLSAQGGPVQS
KSPRFASWDEMNVLAHGLLQLGQGLREHAERTSQLSACGSACQGTGSTDPLAPESRVDPEVLHSLQT
QLKAQNSRIQQLFHKVAQQQRHLEKQHLRIQHLQSQFGLLDHKHDHEVAKPARRKRLPEMAQPVDPAHNVSRLHRL
PRDCQELFQVGERQSGLFEIQPQGSPPFLVNCKMTSDGGWTVIQRRHDGSVDFNRPWEAYKAGFGDPHGEFWLGLEK
VHSITGDRNSRLAVQLRDWDGNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPPSGLSVPFSTWDQDHDLRRD
KNCAKSLSGGWWFGTCSHSNLNGQYFRSI PQQRQKLKKGIFWKWRGRYYPLQATTMLIQPMAAEAAAS

Transmembrane domain.

amino acids 49-69

N-glycosylation site.

amino acids 224-227

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 46-49, 118-121

N-myristoylation sites.

amino acids 50-55, 129-134, 341-346, 357-362

Fibrinogen beta and gamma chains C-terminal domain signature.

amino acids 396-408

Fibrinogen beta and gamma chains.

amino acids 231-447

FIGURE 31

MLLWVILLVLAPVSGQFARTPRPIIIFLQPPWTTVFQGERVLTCKGFRFYSPQTKWYHRYLGKEILRETPDNILEV
QESGEYRCQAQGPLSSPVHLDFSSEMGFPHAAQANVELLGSSDLLT

Signal sequence.
amino acids 1-15

N-myristoylation site.
amino acids 89-94

FIGURE 32

MAGPRPSPWARLLAALISVLSGTLANRCKKAPVKSCTECVRVDKDCAYCTDEMFRDRRCNTQAELLAAGCQRESI
 VVMESFFQITEETQIDTTLRSQMSPOGLRVRLRPGEERHELEVFEPLESPVDDYILMDFNSNSMSDDLDNLKKMGQ
 NLARVLSQLTSQTYTIGFGKFVDKVSPQTDMPKPEKLKEPWNSDPPFSFKNVISLTEDVDEFRNKLQGERISGNLDA
 PEGGFDAILQTAVCTR DIGWRPDSTHLLVFSTESAFHYEADGANLAGIMSRNDERCHLDTGTYTQYRTQDYPSPVP
 TLVRLLAKHNIPIFAVTNYSYYEKHTYFPVSSLGVLQEDSSNIVELLEAFNRIRSNDIRALDSPRGLRTEV
 TSKMFQKTRTGSFIRRGEVGIYQVQLRALEHVGDTHVCQLPEDQKGNIHLKPSFSDGLKMDAGIICDVCTCELQKE
 VRSARCSFNGDFVCQCVCSSEGWSGQTNCSTGSLSDIQPCPLREGEDKPCSGRGEQCQCGHCVCYGEGRYEGQFCYD
 NFQCPRTSGFLCNDRGRCSMGCVCCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGRCHCHQQSLYTDTICEI
 NYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTEECNFVKVMVDELKRAEEVVVRCSFRDEDDCTSYTMEGDGAPG
 PNSTV рр VHKKDCPPGSFWWLIPLLLLLPLLALLLCWKYCACCKACALLPCCNRGHMVGFKEDHYMLRENIMA
 SDHLDTPMLRSGNLKGRDVVRWKVTNNMRQPGFATHAASINPTELV рр GLSRLARLCTENLLKPDTRCAQLRQE
 EENLNEVYRQISGVHKLQQTFRQOPNAGKKQDHITV рр TLMAPRSAKPALLKLTKEQVRAFHDLKVAPGYYLT
 ADQDARGMVEFQEGVELDVDRVPLFIRPDDDEKQLLVEADVPAGTTATLGRRLVNITIIKEQARDVVSFEQPEFSV
 SRGDQVARIPVIRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQPGEAWEKELQVLLLEQEVDSLLRGQVR
 FHVQLSNPKFGAHLGQPHSTIIIIRDПDELDRSFTSQMLSSQPPPQHDLGAPQNPNAKAAGSRKIHFNWLPNGKPM
 GYRVKYWIQGDSESEAHLDSKVPSVELTNLYPYCDYEMKVCAYGAQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVS
 STVTQLSWAEPAETNGEITAYEVCYGLVNDDNRPIGPMKKVLDNPKNRMLLIENLRESQPYRYTVKARNAGWGPE
 REAIINLATQPKRPMSPPIIPDIPIVDAQSGEDYDSFLMYSDDVLRSPSGSQRPSVSDDTEHLVNGRMDFAFPGSTN
 SLHRMTTSAAAYGTHLSPHPHRVLSTSSTLTDYNSLTRSEHSHSTLPRDYSTLTSVSSHDSRLTAGVPDTPTR
 LVFSALGPTSLRVSWQEPRCERPLQGYSVEYQLLNGELHRLNIPNPAQTSVVVEDLLPNHSYVFRVRAQSQEGWGR
 EREGVITIESQVHPQSPCLPLPGSAFTLSTPSAPGPLVFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGPA
 TAFRVDGDSPESLRTV рр PGLSENVРР YKFKVQARTTEGFGPEREGIITIESQDGGFPQLGSRAGLFQHPLQSEYSSIT
 TTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVTQEFSRTLTSGTLSTHMDQQFFQT

Signal sequence.

amino acids 1-23

Transmembrane domain.

amino acids 711-731

N-glycosylation sites.

amino acids 327-330, 491-494, 579-582, 617-620, 695-698, 980-983, 1523-1526

Glycosaminoglycan attachment site.

amino acids 513-516

Tyrosine kinase phosphorylation sites.

amino acids 43-50, 401-408, 674-682, 1596-1604

N-myristoylation sites.

amino acids 24-29, 153-158, 234-239, 420-425, 477-482, 487-492, 560-565, 587-592, 588-593, 782-787, 1031-1036, 1089-1094, 1093-1098, 1383-1388, 1563-1568, 1613-1618, 1614-1619, 1635-1640, 1676-1681, 1719-1724, 1739-1744

Amidation sites.

amino acids 875-878, 974-977

Cell attachment sequence.

amino acids 1003-1006

EGF-like domain cysteine pattern signatures.

amino acids 479-490, 562-573

Integrins beta chain cysteine-rich domain signature.

amino acids 512-525, 590-603

Integrin, beta chain homology.

amino acids 37-455

Calx-beta domain.

amino acids 979-1084

Fibronectin type III domain.

amino acids 1127-1208, 1220-1310, 1458-1542, 1571-1658

FIGURE 33

MTLDRPGEATMLKFTVLLFCIRMSLGMTSIVMDPQPELWIESNYPQAPWENITLWCRSPSRISSKFLLLKDQTQM
TWIRPSHKTFOVSFLIGALTESNAGLYRCCYWKETGWSKPSKVLEAEPGQLPKPIFWIQAETPALPGCNVNILCHG
WLQDLVFMFLKEGYAEPVDYQVPTGTMAIFSIDNLTPEDEGVYICRTHIQMLPTLWSEPSNPLKLVVAGLYPKPTLT
AHPGPIMAPGESLNLRQCQGPIYGMTFALMRVEDLEKSFYHKKTIKNEANFFQSLKIQDTGHYLCFYDASYRGSSL
SDVLKIWVTDTFPKTWLLARPSAVVQMGQNVLRCRGPVGVGLALYKKGEDKPLQFLDATSIDNTSFFLNNVTYS
DTGIYSCHYLLTWKTSIRMPHNTVELMVVDKPKPSLSAWPSTVFKLKGKAITLQCRVSHPVLEFSLEWEERETFQR
FSVNGDFIISNDGKGTGYSCSYRVEHPNMWSHRSEPLKLMGPAGYLWTWNVVLNEAIRLSLIMQLVALLVVLI
RWKCRRRLRIREAWLLGTAQGVVMLFIVTALLCCGLCNGVLIEETEIVMPTPKPELWAETNFPLAPWKNLTLWCRSPS
GSTKEFVLLKDGTVIATPASEQVRRAFPLGALTQSHTGSYHCHSWEEAMVSEPESEALELVGTDILPKPVISASPT
IRGQELQLRCKGWLAGMGFALYKEGEQEPPVQQLGAVGREAFFTIQRMEDKDEGNYSCRTHTEKLPFKWSEPSEPLEL
VIKEMYPKPFKWTASPVVTPGARVTFNCSTPHQHMSFILYKDGSEIASSDRSWASPGASAHHFLIIISVGIGDGGNY
SCRYYDFSIWSEPSDPVELVVTEFYPKPTLLAQPGPVVFPGKSVILRCQGTFOGMRFALLQEGAHVPLQFRSVSGNS
ADFLLHTVGAEDSGNYSCIYYETTMSNRGSYLSMPLMIWVTTDFPKWLFAEPSSVPMQGVTLWCRGPVHGVGYI
LHKEGEATSMQLWGSTSNDGAFPITNISGTSMGRYSCCYHPDWTSSIKIQPSNTLELLVTGLLPKPSLLAQPGPMVA
PGENMTLQCQGELPDSTFVLLKEGAQEPLEQQRPSGYRADFWMPAVRGEDQSIYSCVYLDSTPFAASNHSDSLEIW
VTDKPPKPSLSAWPSTMFKLGKDTLQCRGPLPGVEFVLEHDGEEAPQQFSEDGFVINNVEGKGIGNYSCSYRLQA
YDPDIWSEPSDPLELVGAAGPVAQECTVGNIVRSSLIVVVVVVALGVVLAIEWKKWPRLTRGSETDGRDQTIALEECN
QEGERPGTPANSPSSTSQRISVELPVPI

Signal sequence.

amino acids 1-28

Transmembrane domains.

amino acids 517-537, 555-575, 1261-1281

N-glycosylation sites.

amino acids 53-56, 338-341, 374-377, 381-384, 607-610, 747-750, 798-801,
846-849, 939-942, 986-989, 1027-1030, 1082-1085, 1147-1150, 1223-1226

Tyrosine kinase phosphorylation sites.

amino acids 287-295, 1125-1132

N-myristoylation sites.

amino acids 102-107, 145-150, 195-200, 254-259, 305-310, 336-341, 349-354,
388-393, 480-485, 555-560, 573-578, 656-661, 709-714, 746-751, 792-797, 814-819,
828-833, 840-845, 844-849, 845-850, 897-902, 901-906, 933-938, 938-943, 953-958,
984-989, 1015-1020, 1030-1035, 1130-1135, 1222-1227, 1276-1281, 1315-1320

Immunoglobulin domains.

amino acids 51-108, 145-201, 241-298, 336-394, 434-486, 605-662, 696-752,
792-851, 888-944, 984-1040, 1080-1136, 1176-1228

FIGURE 34

MAPEPAPGRTMVPALVMLGLVAGAHGDSKPVFIKVPEQDTGLSGGVASFVCQATGEPKPRITWMKKGVSSQR
 FEVIEFDDGAGSVLRIQPLRVQRDEAIYECTATNSLGEINTSAKLSVLEEEQLPPGFPSIDMGPQLKVVEKARTATM
 LCAAGGNPDPEISWFKDFLPVDPATSNGRIKQLRSGALQIESSEESDQGKYECVATNSAGTRYSAPANLYVRVRRVA
 PRFSIPSSQEVMPPGSVNLTCVAVGAPMPYVKWMMGAEEELTKEDEMPVGRNVLELSNVVRSANYTCVAISLGMIE
 ATAQVTVKALPKPPIDLVVTETTATSVTLWDGNSEPVYYGIQYRAAGTEGPFQEVGDGVATTRYSIGGLSPFSEY
 AFRVLAVNSIGRGPSEAVRARTGEQAPSSPPRRVQARMLSASTMLVQWEPPPEPNGLVRGYRYYT PDSRRPPNAW
 HKHNTDAGLLTVGSLLPGITYSLRVLAFTAVGDPSPTIQVKTQQGVPAQPADFQAEVESDTRIQLSWLLPPQER
 IIMYELVYWAAEDEDQOHKVTFDPTSSYTLEDLKPDTLYRFQLAARSMDMGVGVFTPTIEARTAQSTPSAPPQKVMCV
 SMGTTVRVSWVPPPADSNGVITQYSVAHEAVDGEDRGRHVVDGISRREHSSWDLVGLEKWTEYRVWVRAHTDVPGP
 PEPPSPVLRVRTDEDVPSGPPRKVEVEPLNSTAVHVYWKLPVFSKQHGQIRGYQVTYVRLENGEPRLPIIQDVMLAEA
 QWRPEESEDYETTISGLTPETTYSVTVAAYTTKGDGARSKPKIVTTGAVPGRPTMMIISTAMNTALLQWHPPKELP
 GELLYGYRLQYCRADEARPNTIDFGKDQHFTVTGLHKGTTYI FRLAAKNRAGLGEFEKEIRTPEDLPSGFQONLHV
 TGLTTSTTELAWDPPVLAERNGRIISYTVVFRDINSQOELQNITTDRFTLTGLKPDTTYDIKVRRAWTSKGSGPLSP
 SIQSRTMPVEQVFAKNFRVAAMKTSVLLSWEVPSYKSAVPFKILYNGQSVEVDGHSMRKLIADLQPNTEYSFVLM
 NRGSSAGGLQHLVSI RTAPDLLPHKPLPASAYIEDGRFDLSMPHVQDPSLVRWFYIVVVPIDRVGGSMLT PRWSTPE
 ELELDELLEAIEQGGEERQRRRQAERLKPYVAAQLDVLPETFTLGDKKNYRGFYNRPLSPDLSYQCFVLASLKEPM
 DQKRYASSPYSDEIVVQVTPAQQQEEPEMLWVTGPVLAVIILIVIAILLFKRKRTHSPSKDEQSIGLKDSLH
 SSDPVEMRLNYQTPGMRDHPIPI TDADNIELRKANDGLKFSQEYESIDPGQQFTWENSLEVNKPKNRYANVIA
 YDHRSVILTSIDGVPGSDYINANYIDGYRKQNAYIATQGPLPETMGDFWRMVWEQRTATVVMTRLEEKSRVKCDQY
 WPARGTETCGLIQVTLLDTVELATYTVRTFALHKGSSEKRELROFQFMAWPDHGVPEYPTPILAFLRRVKACNPLD
 AGPMVVHCSAGVGRGCFIVIDAMLERMKHEKTVDIYGHVTCMRSQRNYMVQTEDQYVFIEALLEATCGHTEVPA
 RNLYAHIQKLGQVPPGESVTAMELEFKLLASSKAHTSRFISANLPCNKFKNRLVNIMPYELTRVCLQPIRGVEGSDY
 INASFLDGYRQQKAYIATQGPLAESTEDFWRMLWEHNSTIIVMLTKLREMGREKCHQYWPAERSARYQYFVVDPMAE
 YNMPQYILREFKVTDARDGSRTIRQFQFTDWPEQGVPKTGEGFIDFIGQVHKTKEQFGQDGPITVHCSAGVGRGTV
 FITLSIVLERMRYEVVDMFQTVKTLRTQRPAMVQTEDQYQLCYRAALEYLGSFDHYAT

Signal sequence.

amino acids 1-29

Transmembrane domain.

amino acids 1262-1282

N-glycosylation sites.

amino acids 117-120, 250-253, 295-298, 721-724, 966-969, 1696-1699, 1731-1734
 cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 71-74, 1286-1289

Tyrosine kinase phosphorylation sites.

amino acids 97-105, 773-780, 1229-1237, 1687-1694

N-myristoylation sites.

amino acids 22-27, 44-49, 47-52, 214-219, 246-251, 305-310, 368-373, 442-447,
 470-475, 481-486, 739-744, 818-823, 881-886, 926-931, 1081-1084, 1301-1306,
 1349-1354, 1362-1367, 1399-1404, 1468-1473, 1551-1556, 1688-1693, 1790-1795,
 1807-1812, 1842-1847, 1847-1852

Amidation site.

amino acids 69-72

Leucine zipper pattern.

amino acids 1262-1283

Myb DNA-binding domain repeat signature 1.

amino acids 1151-1159

Tyrosine specific protein phosphatases active sites.

amino acids 1546-1558, 1837-1849

Immunoglobulin domains.

amino acids 47-109, 149-209, 246-300

Fibronectin type III domains.

amino acids 319-401, 413-500, 512-594, 606-696, 708-809, 821-904, 915-1000

Protein-tyrosine phosphatase homology.

amino acids 1375-1606, 1664-1897

FIGURE 35

MRRAALWLWLCALALSLQLALPQIVATNLPPEDQDGSGDDSDNFSGSGAGALQDITLSQQTSTWKDTQLLTAIPTS
PEPTGLEATAASTSTLPAGEGPKEGEAVVLPEVEPGLTAREQEATPRPRETTQLPTTHQASTTATTAAQEPATSHPH
RDMQPQGHETSTPAGPSQADLHTPHTEDGGPSATERAAEDGASSQLPAAEGSGEQDFTFETSGENTAVVAVEPDRRN
QSPVDQGATGASQGLLDRKEVLGGVIAGGLVGLIFAVCLVGFMLYRMKKDEGSYSLEEPKQANGGAYQKPTKQEEFYA

Signal sequence.

amino acids 1-22

Transmembrane domain.

amino acids 252-272

N-glycosylation site.

amino acids 43-46

Glycosaminoglycan attachment sites.

amino acids 45-48, 47-50

Tyrosine kinase phosphorylation site.

amino acids 279-286

N-myristoylation sites.

amino acids 46-51, 82-87, 183-188, 238-243, 241-246, 254-259, 255-260, 259-264,
263-268

Syndecans signature.

amino acids 276-289

Syndecan domain.

amino acids 3-308

FIGURE 36

MKIFLPVLLAALLGVERASSLMCFSCLNQKSNLYCLKPTICSDQDNYCVTVSASAGIGNLVTFGHSLSKTCSPACPI
PEGVNNGVASMGISSCQSFLCNFSAADGGLRASVTLLGAGLLLSSLPLLRFGP

Signal sequence.

amino acids 1-20

Transmembrane domain.

amino acids 108-128

N-glycosylation site.

amino acids 99-102

N-myristoylation sites.

amino acids 14-19, 58-63, 80-85, 89-94, 105-110, 106-111, 117-122

u-PAR/Ly-6 domain.

amino acids 21-100

FIGURE 37

MDGKKCSVWMFLPLVFTLFTSAGLWIVYFIAVEDDKILPLNSAERKPGVKHAPYISIAGDDPPASCVFSQVMNMAAF
LALVVAVLRFIQLKPKVLNPWLNISGLVALCLASFGMTLLGNFQLTNDEEIHNVGTSLTFGFTLTCWIQAALTAKV
NIKNEGRRGIPRVIILSASITLCVVLYFILMAQSIHMYAARVQWGLVMCFLSYFGTFAVEFRHYRYEIVCSEYQENF
LSFSESLSEASEYQTDQV

Transmembrane domains.

amino acids 7-27, 65-85, 97-117, 131-151, 166-186

N-glycosylation site.

amino acids 100-103

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 4-7

N-myristoylation sites.

amino acids 132-137, 140-145, 199-201

Amidation sites.

amino acids 2-5, 159-162

FIGURE 38

MELALLCGLVVMAGVIPIQGGILNLNKMVKQVTGKMPILSYWPYGC~~H~~GLGGRGQPKDATDWCCQTHDCCYDHLKTQ
GCGIYKDYYRYNFSQGNIHCSDKG~~S~~WCEQQLCACDKEVA~~F~~CLKRNLDTYQKRLRFYWRPHCRGQT~~PG~~C

Signal sequence.
amino acids 1-17

N-glycosylation site.
amino acids 89-92

N-myristoylation sites.
amino acids 20-25, 45-50, 93-98, 140-145

Phospholipase A2 histidine active site.
amino acids 63-70

Phospholipase A2 aspartic acid active site.
amino acids 108-118

Phospholipase A2 homology.
amino acids 21-145

FIGURE 39

MYIRVSYDTKPDSSLHLMVKDWQLELPKLLISVHGLQNFEMQPKLKQVFGKGLIKAAMTTGAWIFTGGVSTGVISH
 VGDALKDHSSSKSRGRVCAIGIAPWGIVENKEVLGVKDVTRYQTMSNPLSKLSVLNNNSHTHFILADNGTLGKYGAEV
 KLRRLLEKHISLQKINTRLGQGVPLVGLVVEGGPNVVSIVLEYLQEEPPIPVVICDGSGRASDILSFAHKYCEEGGI
 INESLREQLLVTIQKTFNYNKAQSHQLFAIIMECMKKELVTVFRMGSEGQQDIEMAILTALLKGTNVSAFDQSLA
 LAWNRVDIARSQIFVFGPHWTPGLSLAPPDKATEKEKKPPMATTKGGRGKGKKGKVKEEVEETDPRKIELL
 NWVNVALEQAMLDALVLDRVDFVKLLIENGVNMQHFLTIPRLLELYNTRLGPPNTHLLVRDVKSNLPPDYHISLID
 IGLVLEYLMGGAYRCNYTRKNFRTLYNNLFGPKRPKALKLGMEDDEPPAKGKKKKKKKEEEIDIDVDDPAVSRFQ
 YPFHELMVWAFLMKRQKMAFLWQRGEESMAKALVACKLYKAMAHESESSDLVDDISQDLDDNSKDFGQLALELLDQ
 SYKHDCEAMKLLTYELKNWSNSTCLKLAVAAKHRDFIAHTCSQMLLTDMWMGRRLMRKNPGLKVIDMILLPPTILF
 LEFRTYDDFSYQTSKENEDGKEEEENTDANADAGSRKGDEENEHKKQRSIPIGTKICEFYNAPIVKFWFTISYLG
 YLLLNFNYVILVRMDGWPSLQEWINVISYIVSLALEKIREILMSEPGKLSQKIKVWLQEWNIITDLVAISTFMIGAILR
 LQNQPYMGYGRVIYCVDIIFWYIRVLDIFGVNKYLGPYVMMIGKMMIDMLYFVVIMLVVLMSFGVARQAILHPEEK
 SWKLARNIFYMPYWMIFYGEVFADQIDLYAMEINPPCGENLYDEEGKRLPPCIPGAWLTPALMACYLLVANILLVNLL
 IAVFNNTFFEVKSISNQVWKFQRYQLIMTFHDRPVLPMMIILSHIYIIIMRLSGRCRKREGDQEERDRGLKLF
 DEELKRLHEFEEQCVQEHFREKEDEQQSSSDERIRVTSERVENMSMRLEEINERETFMKTSQTVDLRLAQLEELSN
 RMVNALENLAGIDRSRDLIQARSRASSECEATYLLRQSSINSADGYSLYRYHFNGEELLFEDTLSLSTSPGTGVRKKT
 SFRIKEEKDVKTHLVECPQNSLHLDSLGTSTSATPDGSHLAVDDLKNAEESKLGPDIGISKEDDERQTDKKEETISP
 SLNKTVDIHGQDKSDVQNTQLTVETTNIEGTISYPLEETKITRYFPDETINACKTMKSRSFVYSRGRKLVGGVNQDV
 EYSSITDQQLTTEWQCQVQKIRSHSTDIPYIVSEAAVQAEQKEQFADMQDEHHVAEAIPRIPLSLTITDRNGMEN
 LLSVKPDQTLGFPSSLRSLKSLHGHPRNVKSIQGKLDLRSGHASSVSSLVIVSGMTAEKKVKKEKASTETEC

Transmembrane domains.

amino acids 759-779, 828-848, 857-877, 893-913, 976-996, 992-1012, 1031-1051

N-glycosylation sites.

amino acids 133-136, 144-147, 233-236, 298-301, 478-481, 601-604, 635-638, 638-641, 830-833, 1006-1009, 1121-1124, 1312-1315

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 1228-1231

Tyrosine kinase phosphorylation sites.

amino acids 697-704, 891-898

N-myristoylation sites.

amino acids 35-40, 53-58, 68-73, 69-74, 102-107, 211-216, 229-234, 296-301, 473-478, 728-733, 747-752, 1166-1171, 1259-1264, 1268-1273, 1319-1324

Amidation sites.

amino acids 362-365, 513-516, 968-971, 1374-1377

Ion transport protein homology.

amino acids 789-1005

FIGURE 40

MSTEKVDQKEEAGEKEVCGDQIKGPDKEEEPPAASHGQGWRPGGRAARNARPEPGARHPALPAMVNNDPPVPALLWA
QEVGQVLAGRARRLLLQFGVLFCTILLWWVSFVLYGSFYYSYMPVTSHLSPVFYYRTDCDSSTTSLCSFPVANVS
LTKGGGRDRVLMYQGPYRVTLELELPESPVNQDLGMFLVTISCYTRGGRIISTSSRSVMLHYRSDLQMLDTLVFSSL
LLFGFAEQKQLLEVELYADYRENSYVPTTGAIIEIHSKRQLYGAYLRIAHFTGLRYLLYNFPMTCAFIGVASNFT
FLSVIVLFSYMQWVWGGIWPRHRFSLQVNIRKDNRKEVQRRISAHQPGPEGQEESTPQSDVTEDGESPEDPSGTE
GQLSEEKPDQQPLSGEEELEPEASDGGSWEDAALLTEANLPAPAPASASAPVLETLGSSPAGGALARQRPTCSSS

Transmembrane domains.

amino acids 94-114, 134-154, 292-312, 308-328

N-glycosylation sites.

amino acids 152-155, 306-309

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 350-353

Tyrosine kinase phosphorylation site.

amino acids 240-248

N-myristoylation sites.

amino acids 44-49, 81-86, 96-101, 302-307, 361-366

Leucine zipper patterns.

amino acids 84-105, 91-112

FIGURE 41

MGPPPLPLLLLLLPPRVLPAAPSSVPRGRQLPGRGLCILEEGLCASEACVNDGVFGRCQKVPAMDFYRYEVSPV
 ALQRLRVALQKLSGTGFTWQDDYTQYVMDQELADLPKTYLRRPEASSPARPSKHSVGSERYSREGGAALANALRRH
 LPFLEALSQAPASDVLARTHTAQDRPPAEGDDRFSESISILTYVAHTSALTYPPGPRTQLREDLLPRTLQGQPDELSP
 KVDSGVDRHHLMMAALSAYAAQRPPAPPGEGLSLEPQYLLRAPSMPRPLLAPAAPQKWPSPLGDSEDPSSTGDGARIH
 TLLKDLQRQPAEVRLSGLELDGMAELMAGLMQGVDHGVARGSPGRAALGESGEQADGPKATLRGDSFPDDGVQDDD
 DRLYQEvhrlSATLGGLLQDHGSRLPGALPFARPLDMERKKSEHPESSLSEEETAGVENVKSQTYSKDLLGQQPH
 SEPGAAAFGELQNQMPGPSKEEQLSPAGAQEALSDGLQLEVQPSEEARGYIVTDRDPLRPEEGRRLVEDVARLLQV
 PSSAFADVEVLGPATFKVSANVNVTTEDEVEKATVDNKDKLEETSGLKILQTGVGSKSKLKFLPPQAEQEDSTKFI
 ALTLVSLACILGVLLASGLIYCLRHSSQHRLKEKLSGLGGDPGADATAAYQELCRORMATRPPDRPEGPHTSRISSV
 SSQFSDGPIPSPSARSSASSWSEEPVQSNMDISTGHMILSYMEDHLKNKNRLEKEWEALCAYQAEPNSSFVAQREEN
 VPKNRSLAVLTYDHSRVLLKAENSHSHSDYINASPIMDHDPRNPAYIATQGPLPATVADFQMVWESGCVVIVMLTP
 LAENGVRQCYHYWPDEGSNLYHIYEVNLVSEHIWCEDFLVRSFYLNLTNETRTVTQFHFLSWYDRGVPSRSSRL
 DFRRKVNCYRGRSCPIIVHCSGDGAGRSGTYVLI DMVLNKMAKGAKIEDIAATLEHLRDQPGMVQTKEQFEFALTA
 VAEEVNAILKALPQ

Signal sequence.

amino acids 1-21

Transmembrane domain.

amino acids 616-636

N-glycosylation sites.

amino acids 564-567, 760-763, 774-777, 898-901

Glycosaminoglycan attachment sites.

amino acids 90-93, 652-655

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 137-140, 425-428

N-myristoylation sites.

amino acids 44-49, 47-52, 91-96, 144-149, 338-343, 342-347, 346-351, 490-495,
 628-633, 634-639, 655-660, 659-664, 852-857, 915-920, 948-953, 987-992

Amidation site.

amino acids 525-528

Cell attachment sequence.

amino acids 372-374

Tyrosine specific protein phosphatases active site.

amino acids 943-955

Protein-tyrosine phosphatase.

amino acids 770-1004

FIGURE 42

MQPPPSLCRALVALVLACGLSRIWGEERGFPPDRATPLLQTAEIMTPPTKTLWPKGNSASLARSLAPAEVPKGDRTAGSPPRTISPPPCQGPIEIKETFKYINTVVSCLVFVLGIIGNSTLLRIIYKNKCMRNGPNILIASLALGDLHIVIDIPINVYKLLAEDWPFGAEMCKLVPFIQKASVGITVSLCALSIDRYRAVASWSRIKGIGVPKWTAVEIVLIWVVSVVLAVPEAIGFDIITMDYKGSYLRICLLHPVQKTAFMQFYKTAKDWWLF SFYFCLPLAITAFFYTLMTCEMLRKYSGMQIALNDHLKQRREVAKTVFCLVLVFALCWPLHLSRILKLTLYNQNDPNRCELLSFLLVLDYIGINMASLNSCINPIALYLVSKRFKNCFKSCLCCWCQSFEEKQSLEEKQSLKFKANDHGYDNFRSSNKYSSS

Transmembrane domains.

amino acids 1-19, 101-121, 137-157, 177-197, 216-236, 275-295, 323-343, 362-382

N-glycosylation sites.

amino acids 59-62, 119-122

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 302-305

Tyrosine kinase phosphorylation site.

amino acids 424-430

N-myristoylation sites.

amino acids 57-62, 115-120, 170-175, 306-311, 371-376

7 transmembrane receptor homology.

amino acids 118-386